AMERICAN Scientist

May–June 2011

www.americanscientist.org

SIGMA XI

THE SCIENTIFIC RESEARCH SOCIETY

AMERICAN
Scientist

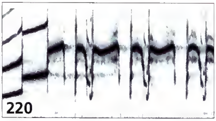
Volume 99 • Number 3 • May–June 2011

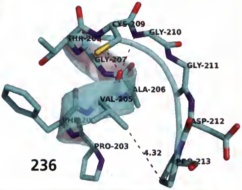
Departments

Scientists' Bookshelf

From Sigma Xi

Feature Articles





The Cover

Female loons typically lay two eggs, a day apart. By the time the second hatches, the first chick is usually already on the water; it is rare to see two chicks in the nest at the same time. If food is plentiful and luck is running with the family, the adults may raise both chicks but often only the dominant chick survives, generally the older one. The presence of two chicks with the loon on the cover means that they have already passed the riskiest part of reproduction; most predation occurs at the egg stage. Yet threats remain from snapping turtles, big muskies, eagles and other predators. In "Marking Loons, Making Progress" (pages 220–227), Walter Piper, Jay Mager and Charles Walcott present the arresting results of nearly two decades of studying banded loons in the wild. Their work retires misconceptions, raises new research questions and sharpens our picture of this symbol of the Northern wild. (Photograph by Tom Walker, Visuals Unlimited, Inc.)

OUR WORLD-UNDERWATER SCHOLARSHIP SOCIETY
Sponsored by Rolex

INVESTING IN THE FUTURE OF DIVING

For over 35 years, a diverse group of divers, marine scientists and underwater explorers known as the Our World-Underwater Scholarship Society has been engaged in a unique mission: to identify and develop the next generation of dive leaders and underwater researchers.

Each spring at the Explorers Club in New York City, the Scholarship Society makes a presentation of the prestigious Rolex Scholarships to three outstanding young students with a demonstrated interest in underwater-related careers. During their scholarship year, the recipients have the opportunity to travel internationally and learn from the world's leading experts in conservation, underwater photography, maritime archaeology, marine engineering and other related fields.

Additionally, the Scholarship Society offers experience-based summer internships with organizations including the National Park Service, the American Academy of Underwater Sciences (AAUS), Divers Alert Network (DAN), the Reef Environmental Education Foundation (REEF) and Sport Diver magazine. These internships are open to students with specialized interests and backgrounds.

"Our goal is to foster the development of future leaders of the underwater environment," says George Wozencraft, president of the Scholarship Society. Applicants for the scholarships must be between the ages of 21 and 26, have high academic standing and achieved Rescue Diver or equivalent certification. The internships are primarily directed at university undergraduates and recent graduates.

HOW YOU CAN GET INVOLVED

The Scholarship Society invites you to nominate students, provide opportunities for scholars and interns or simply contribute to help support the organization's educational mission.

To learn more, read scholar and intern journals and download applications, go to owuscholarship.org. Deadline for 2012 North American and European Scholarships is December 31, 2011. Deadline for 2012 internships and Australasian Scholarship is January 31, 2012.

OYSTER PERPETUAL SUBMARINER

NORTH AMERICA

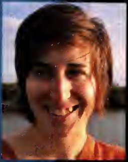
CHRISTIAN CLARK, 23, has felt connected to the water since he was a child. His family lived in Brazil, Canada, and Georgia, never too far from water and boats. Christian learned to dive in Hawaii at age 13. Now a recent graduate of the University of Hawaii with a BS in Global Environmental Science, Christian is a Divemaster with over 330 dives. As a research assistant in the Institute of Marine Biology's Shark Lab, he studied sharks and apex predators around the main and northwest Hawaiian Islands. He joined three research expeditions to Antarctica observing climate change effects on benthic ecology. Christian sees documentary films as effective tools for communicating environmental issues and solutions to the public. He hopes the scholarship will help point him to a career that would combine his passions for marine science, research diving and photography.

EUROPE

ERIN McFADDEN, 21, was raised in the Scottish coastal town of Dunbar and has always found the sea a source of fascination and contentment. Her first dive was in the North Sea at 17. By 18 Erin was a PADI Divemaster, completing her training during a marine conservation expedition to the Bahamas. Now with over 300 dives her experiences have included Divemaster work in Europe and Africa, and scientific diving internships in South Africa and Mozambique. Currently, a final year Marine Biology student at the University of St Andrews in Scotland, Erin believes her studies and dive experiences have helped her appreciate the delicate intricacy of the sea and the need to develop globally effective and sustainable measures of marine conservation and resource management. She views the prospect of a year spent getting to the heart of the diving and underwater world as an unparalleled opportunity.

AUSTRALASIA

ANTHEA IBELL, 22, completed her PADI Open Water Instructor Diploma at Dive Otago, Dunedin in 2008. Since then she has been instructing part time while studying to achieve a Bachelor of Science degree in Zoology at Otago University, with a Minor in Archaeology. Anthea has a passion for marine science and conservation and is interested in species distribution, physiology, behavior and population studies. She has travelled extensively around North and South America which taught her valuable skills and introduced her to different cultures and world views. The Our World-Underwater Scholarship Society's Rolex Scholarship will allow her to bring all her interests together through marine research, SCUBA diving, photography and travel. She aims to contribute to conservation work around the world and increase public awareness about threats facing the oceans today.



OUR WORLD - UNDERWATER SCHOLARSHIP SOCIETY®

ROLEX

When Less Is More

As I write this the troubles at Fukushima Daiichi are 13 days old and far from over. Just as the world was beginning to recover from the shock of the Deepwater Horizon oil spill, and turn a more favorable gaze on nuclear power, the apple cart topples again. Germany appears ready to abandon its nuclear ambitions, and others are sure to follow.

Sad as these events are and will continue to be, they make Vaclav Smil's plaint all the more timely. In "Global Energy: The Latest Infatuations" (pages 212–219) he takes western nations—and the United States and Canada in particular—to task for their continual reliance on some ever-shiftinging technological panacea to deal with energy problems. Particularly damning, even among richer nations, the United States uses twice as much energy per capita to achieve a very similar standard of living. In his usual systematic fashion, the University of Manitoba professor applies a heavy dose of numerical reality to the prospects for nuclear-, solar- or wind-derived electricity allowing this trend to continue. Then he puts the clean-coal crowd away with an analysis of the practicality of carbon capture and sequestration. In end, his answer is prudent and simple: Use less.

You may have noticed that the image at upper left has changed somewhat. Our ever-clever assistant art director has superimposed me on one of my own photographs. Obviously, the combo is sheer fiction; when I made the background photograph, I was behind the camera I'm holding in the photograph (self-referential camera?). I have a hunch that some of the geologists among you will recognize the locale, but here are a few hints: The red formation exposed at my left elbow is the Triassic Chugwater Group. For those of you not so geologically inclined, you might also have passed by on the way to a place central to the early women's suffrage movement. You would have "passed" it about 10 miles west of the roadside pullout where the Fujichrome was exposed.

Finally, I must relay sad news. George Bugliarello, one of *American Scientist*'s staunchest proponents and allies for the past several decades, passed away on February 18. You may be familiar with part of George's legacy as the George Bugliarello Prize, an award given every two years to the *American Scientist* article that best integrates "technology, human society, our biological needs and the needs of other life on our planet...." Fred Gould, Krisztian Magori and Yunxin Huang were the first recipients for their article on genetic control of mosquito-borne diseases, and last year, David T. Wong received the award for his article "Salivary Diagnostics."

Besides being a great friend of the magazine, George was an engineer, president of Sigma Xi, president of Polytechnic Institute in New York, a member of the National Academy of Engineering, a founding fellow of the American Institute of Medical and Biological Engineering, and a true gentleman. He is missed.—*David Schoonmaker*

AMERICAN
Scientist
www.americanscientist.org

VOLUME 99, NUMBER 3

EDITORIAL AND SUBSCRIPTION CORRESPONDENCE
American Scientist
P.O. Box 13975
Research Triangle Park, NC 27709
919-549-0097 • 919-549-0090 fax
editors@amscionline.org • subs@amsci.org

American Scientist gratefully acknowledges support for "Engineering" through the Leroy Record Fund.

Sigma Xi, The Scientific Research Society was founded in 1886 as an honor society for scientists and engineers. The goals of the Society are to foster interaction among science, technology and society; to encourage appreciation and support of original work in science and technology; and to honor scientific research accomplishments.

Printed in USA

Travel to the Stars

To the Editors:

Regarding the very interesting article "Ancestors of Apollo" by Dennis Danielson (March–April), I would like to note that Copernicus had a little help coming up with his heliocentric theory. The Greek philosopher Aristarchus of Samos had the same idea around 200 B.C., some 17 centuries before Copernicus. Several historians have noted that Copernicus was aware of Aristarchus's heliocentric concept. So, in a sense, the ancestors of Apollo go back at least 2,200 years.

Dates aside, a key point in the article is that scientists need to struggle onward even if it takes centuries for their work to come to fruition. That idea needs to be emphasized again and again. As an example of really speculative science today, people are thinking about how to travel to the stars. In 1994, I co-led a NASA-sponsored workshop on advanced physics concepts that might provide insight into faster-than-light (FTL) travel. Although no obvious methods emerged during that meeting, physicists continue to study the theoretical basis for FTL movement that could make interstellar travel conceivable. Just as ancient astronomers dealt with a geocentric universe and impenetrable crystalline spheres, today we deal with theories that say travel to the stars is impossible. What our age needs is another Copernicus (or Aristarchus)!

Gary L. Bennett
Boise, ID

Level-headed Peer Review

To the Editors:

I commend you for publishing bioethicist David B. Resnik's essay "A Troubled Tradition" (January–February). During my career as an agronomist, I have reviewed many journal articles, book chapters, research proposals, etc. When I thought papers merited publishing, I usually remained anonymous and tried to help authors with whatever suggestions I could. But a few papers simply did not measure up. When that happened, I sent copies of my rejection letters to the senior authors, hoping that would help them put my position in perspective. That was fine with some editors. Others, insisting on anonymity, stopped sending me papers. That was helpful for my busy schedule.

More needs to be said about the consequences of the peer-review process. Consider young scientists. After working several years, they can have a parental attitude toward their early research. If that is rejected during a blind peer-review process, scientists can feel as if someone wounded their children. On top of that, the professional stakes are high. Employers frequently judge scientists by the number of publications they achieve. With so much in play, scientific societies must continuously audit the performance of their editors.

I always told graduate students and postdoctoral scientists that the peer-review process can be capricious and that if their paper was accepted, they should not get a big head. Along the same lines, if it was rejected, I told them not to let themselves be pulled down. Instead, I recommended that we work together to see what we could learn from the rejection.

Donald N. Baker
Starkville, MS

To the Editors:

David Resnik has discovered that editors and reviewers of scientific publications are all too human. I have done reviewing and editing in several disciplines and I can vouch that what he

American Scientist (ISSN 0003-0996) is published bimonthly by Sigma Xi, The Scientific Research Society, P.O. Box 13975, Research Triangle Park, NC 27709 (919-549-0097). Newsstand single copy $4.95. Back issues $6.95 per copy for 1st class mailing. U.S. subscriptions: one year $28, two years $50, three years $70. Canadian subscriptions: one year $36; other foreign subscriptions: one year $43. U.S. institutional rate: $70, Canadian $78; other foreign $85. Second-class postage paid at Durham, NC, and additional mailing office. Postmaster: Send change of address form 3579 to Sigma Xi, P.O. Box 13975, Research Triangle Park, NC 27709. Canadian publications mail agreement no. 40040263. Return undeliverable Canadian addresses to P. O. Box 503, RPO West Beaver Creek, Richmond Hill, Ontario L4B 4R6.

says is correct. Neither reviewers nor editors are as objective or fair as is naively assumed by many scientific outsiders and a few insiders.

There is no system that is going to solve this problem. The Founding Fathers wrote the U.S. Constitution and yet politicians have always figured out ways to get around it. So think about that when you vote for the officers of your scientific societies. Officers should think about it when they appoint editors. Editors should think about it when they recruit reviewers. And we all should think about it when we vote for our elected representatives.

In certain physical sciences, the publication of something incorrect or even fraudulent may have little consequence, if the topic is sufficiently trivial. In the medical sciences, however, people may start dying.

Foster Morrison
North Potomac, MD

The Perfect Dome

To the Editors:

Henry Petroski's article "Arches and Domes" (March–April) reminded me that the earliest masonry domes were probably snow houses or igloos, rather than highly decorated houses of worship. The igloo is not a hemisphere, which would collapse as the snow compresses and the sides bulge out. Instead, igloos follow a more complex curve called a catenary. Catenary, which derives from the Latin word for chain, describes the curve taken by a hanging chain that is held at both ends. As it compresses, an igloo retains its catenary shape and simply becomes shorter. When it gets too short, it is time to build a new igloo.

The igloo in the photo (*below*) was built in 1970 as a demonstration by Inuit of Pond Inlet, Nunavut, Canada. Once it was complete, I tested its strength with a device that measures the snow's cohesion and internal friction. Based on the results, I calculated that an igloo could have a maximum diameter of 10 meters. An igloo of approximately that dimension, built for partying, has indeed been reported in the literature.

An igloo's catenary shape contrasts with that of a Gothic arch. The Gothic arch has a circular cross section with the top cut out and the sides pushed together. That was not a good move in terms of stability, because it is the part that is cut out that would be stable without any extras like gargoyles or flying buttresses to transfer stresses to adjacent rows of columns. If cathedral designers had studied a length of chain held by its ends, they might have stumbled upon a better way to shape an arch. A few well-known structures, such as the Gateway Arch in St. Louis and the domes of St. Paul's Cathedral in London and the U.S. Capitol, do approximate catenaries. But it is the Inuit who capitalized on the catenary without embellishments.

Richard L. Handy
Iowa State University

Cave Controversy

To the Editors:

Nine biospeleologists from Brazil, Slovenia, the United Kingdom and the United States find that Aldemaro Romero's "The Evolution of Cave Life" (March–April 2011) misrepresents our current and past views. Thomas C. Barr, Kenneth Christiansen, Annette Summers Engel, John Holsinger, Matthew Niemiller, Graham Proudlove, Boris Sket, Eleonora Trajano and I work in different aspects of biospeleology on different organisms and have written hundreds of peer-reviewed papers over the past 50 years. On the basis of our wide experience and research, we find that Romero gives, at worst, a biased view and at best a minority view of important aspects of biospeleology. We fear that the naive reader will be misled by his erroneous subtitle, "New concepts are challenging conventional ideas about life underground." We find:

- Errors of terminology (e.g., troglobitic vs. troglomorphic) and biology (e.g., one-third of hypogean fish are fully eyed and pigmented).
- Misrepresentations of the data and conclusions of our published studies. Not one of us argues that troglomorphic characters necessarily develop in parallel. Nor do we argue that caves are so constant that no ecological fluctuations take place.
- Incorrect characterization of our views as unitary and neo-Lamarckian. We argue concepts all the time, and we are all strong Darwinists.
- "Straw-man" arguments, especially about pre-adaptation and archetypes. None of us has written that all troglobites had ancestors that were preadapted, and not one of us believes in an archetypical hypogean species.
- Exclusive emphasis on his single hypothesis of the centrality of phenotypic plasticity for the colonization of caves. We explore all of the multiple hypotheses for colonization or isolation in caves.
- Exclusive emphasis on only one study system, that of the cave and surface Mexican *Astyanax* fish, for which only some populations support some of his views.

Romero's article is not as strident as his 2009 book, *Cave Biology: Life in Darkness*. But our general areas of concern and disagreement with his article are the same as those several of us presented in reviews of his book. These include my review in the spring 2011 issue of *Bioscience* and Engel's review in the October 2010 issue of *Integrative and Comparative Biology*.

Thomas Poulson
Jupiter, FL

To the Editors:

I felt dissonance reading Aldemaro Romero's cave article. Although a focus on his own research was understandable, his general approach and lack of reference to other pertinent work implied that biospeleology was in disarray, which it is not.

For example, Romero used definitions for troglobite and troglophile that were based on troglomorphic characteristics, troglobites being very troglomorphic and troglophiles being less so. More customary definitions are based on how strongly troglobites and troglophiles are confined to caves—troglobites being obligate cave dwellers, but troglophiles being able to live both in and out of caves. Troglophiles often show no apparent troglomorphisms

Illustration Credits

Computing Science
Pages 192–194 — Brian Hayes

Marginalia
Page 201 — Tom Dunne

Ethics
Pages 205, 206 — Tom Dunne

Global Energy: The Latest Infatuations
Figures 2–4, 7 — Tom Dunne

Marking Loons, Making Progress
Figures 3, 5, 7, 9–11 — Tom Dunne

Pliocene Climate Lessons
Figures 2, 4–8 — Barbara Aulicino

Porphyrins: One Ring In the Colors of Life
Figures 2, 5 *(right)* — Barbara Aulicino

due to ongoing genetic exchange with epigean relatives. Troglobites, which lack such exchange, tend to show considerable troglomorphism, and those that have been isolated in caves the longest show the most. Throughout his article, Romero neglected to emphasize the importance of isolation to the development of troglomorphisms, and the definitions he chose for troglobite and troglophile also ignored the concept of isolation.

As another example, Romero discounted the phenomenon of entrapment as a likely step toward living in a cave, but failed to explain the well-grounded example of how Ice Age climate change could do just that. For example, if the surface environment became too dry for a particular troglophile, that troglophile would be trapped in its damp cave, epigean relatives would depart or die due to the dry conditions, and gene flow from the surface would stop. The stage would then be set for evolution toward greater cave adaptation. This loss of gene flow from surface relatives answers a question Romero posed about why some hypogean species undergo major phenotypic changes while others remain similar to their epigean ancestors.

Although this article presented many interesting observations about cave life, it failed to integrate them into our broader understanding of how cave life evolved, particularly with regard to the role played by isolation, the length of that isolation and our understanding gained through comparative biogeography.

Norman W. Youngsteadt
Springfield, MO

To the Editor:

Aldemaro Romero's excellent and informative article on cave animals not only makes good points about the adaptive genetics of cave life but also illustrates how organisms of choice can affect conclusions. This speleobiologist, who focuses on troglobiotic terrestrial arthropods in temperate climates, sees some matters differently from one who studies aquatic vertebrates in mostly tropical environments. As an example, Romero does not mention the climatic relict hypothesis of cave invasion in answering the question of how animals get into caves to begin with. In this idea, the ancestors of troglobionts are not cave invaders by choice (as Romero's fish might be) but are marooned in caves when the climate outside changes drastically and surface populations become locally extinct. Similarly, since most cave arthropods are members of taxa which inhabit cave-like places such as the soil and leaf litter, preadaptation to those habitats very likely plays a role in their success in caves. In order to fully understand the evolution of cave animals it is necessary to examine a wide spectrum of taxa.

William Shear
Hampden-Sydney College

Dr. Romero responds:

The letter from Thomas Poulson et al. makes a number of claims that are not substantiated by reading my article. The researchers claim that I misrepresent their views even though they are not even mentioned in the article. They also claim that I emphasize a single hypothesis or study system when describing certain phenomena of cave life. That was precisely the point of the article: to show that there are data challenging the orthodoxy carried forward by some people for some time. To examine the full range of hypotheses and ideas, readers may wish to look at my 2009 book.

From reading their letter it is clear that the main contention of these authors is that I do not adhere to the right orthodoxy. As I analyzed in great detail in chapter 1 of my book, supported by extensive documentation, biospeleology is a field that has been plagued with intellectual inertia over time. If the history of science teaches us anything, it is that change is the only constant and that change must be generated by data, such as those I present in my book and article.

In response to Norman Youngsteadt: My book cites the conventional hypotheses about cave colonization. On the issue of definitions I have purposely moved away from the typological approach to the classification of cave organisms because, like many modern biologists, I do not believe that species are members of a periodic table like the elements of chemistry. Data cited in my book strongly suggest that there is simply too much flux in their ecological and morphological characteristics to believe that all cave populations can be boxed into predetermined types. Regarding isolation, Youngsteadt's reference to the Ice Age is not an example but a hypothesis that has not been corroborated by experimental data. The hypothesis is disqualified as a valid generalization to explain cave colonization if we consider that the vast majority of cave biodiversity is found in the tropics and not in temperate environments (where most cave biologists have traditionally done their studies), and that a large portion of cave populations that display characteristics typically associated with life in darkness date back well before any particular Ice Age.

Shear makes a good point that terrestrial and aquatic organisms will have different restrictions when it comes to dispersal in caves, something I mention in my book. When I refer there to the climatic relict hypothesis, I note that so far the idea lacks empirical evidence; further, its assumptions cannot be generalized to all cave organisms, the vast majority of which are found in the tropics. As I also show in my book, available data do not support the idea of preadaptation as a general explanation for successful colonization of caves. Other characteristics of an organism such as being ecological generalists, probably combined with genetic heterozygocity, likely play an important role. That is also noted in my book, which includes a survey of all taxa represented in caves, from bacteria to mammals. As Theodosius Dobzhansky pointed out, evolution is opportunistic. Cave biotas are a good example of that.

How to Write to *American Scientist*

Brief letters commenting on articles that have appeared in the magazine are welcomed. The editors reserve the right to edit submissions. Please include a fax number or e-mail address if possible. Address: Letters to the Editors, *American Scientist*, P.O. Box 13975, Research Triangle Park, NC 27709 or editors@amscionline.org.

Errata

Owing to a typographical error in "Arches and Domes" (March–April), the photo caption for the Hagia Sophia indicates that Christians and Muslims fought over the structure in the 5th century rather than the 15th.

In "The Evolution of Cave Life" (March–April) Figure 8 shows a skeleton on a cave floor labeled a swiftlet. It is actually a bat.

The Man Behind the Curtain

Tony Rothman

"I want to get down to the basics. I want to learn the fundamentals. I want to understand the laws that govern the behavior of the universe." Thousands of admissions officers and physics department chairs have smiled over such words set down by aspiring physicists in their college-application essays, and that is hardly surprising, for every future physicist writes that essay, articulating the sentiments of all of us who choose physics as a career: to touch the fundamentals, to learn how the universe operates.

It is also the view the field holds of itself and the way physics is taught: Physics is the most fundamental of the natural sciences; it explains Nature at its deepest level; the edifice it strives to construct is all-encompassing, free of internal contradictions, conceptually compelling and—above all—beautiful. The range of phenomena physics has explained is more than impressive; it underlies the whole of modern civilization. Nevertheless, as a physicist travels along his (in this case) career, the hairline cracks in the edifice become more apparent, as does the dirt swept under the rug, the fudges and the wholesale swindles, with the disconcerting result that the totality occasionally appears more like Bruegel's Tower of Babel as dreamt by a modern slumlord, a ramshackle structure of compartmentalized models soldered together into a skewed heap of explanations as the whole jury-rigged monstrosity tumbles skyward.

Of course many grand issues remain unresolved at the frontiers of physics: What is the origin of inertia? Are there extra dimensions? Can a Theory of Everything exist? But even at the undergraduate level, far back from the front lines, deep holes exist; yet the subject is presented as one of completeness while the holes—let us say abysses—are planked over in order to camouflage the danger. It seems to me that such an approach is both intellectually dishonest and fails to stimulate the habits of inquiry and skepticism that science is meant to engender.

> *Physics is not always the seamless subject that it pretends to be*

In the first week or two of any freshman physics course, students are exposed to the force of friction. They learn that friction impedes the motion of objects and that it is caused by the microscope interaction of the two surfaces sliding past one another. It all seems quite plausible, even obvious, yet regardless of any high falutin' modeling, with molecular mountain ranges resisting each other's passage or running-shoe soles binding to tracks, friction produces heat and hence an increase in entropy. It thus distinguishes past from future. The increase in entropy—the second law of thermodynamics—is the only law of Nature that makes this fundamental distinction. Newton's laws, those of electrodynamics, relativity ... all are *reversible*: None care whether the universal clock runs forward or backward. If Newton's laws are at the bottom of everything, then one should be able to derive the second law of thermodynamics from Newtonian mechanics, but this has never been satisfactorily accomplished and the incompatibility of the irreversible second law with the other fundamental theories remains perhaps the greatest paradox in all physics. It is blatantly dropped into the first days of a freshman course and the textbook authors bat not an eyelash.

To a physicist, moreover, the material world is divided into billiard balls and springs. An ideal spring oscillates forever, but anyone who has ever watched a real-world spring knows that forever usually lasts just a few seconds. We account for this mathematically by inserting a frictional term into the spring equation and the fix accords well with observations. But the insertion is completely *ad hoc*, adjustable by hand, and to claim that such a fudge somehow explains the behavior of springs is simple vanity.

Not Exactly

Perhaps our complacency is due to the fact that we have written down a plausible equation. Physicists have long believed that mathematics is the Rosetta Stone for unlocking the secrets of Nature and since a famous 1960 essay by Eugene Wigner entitled "The Unreasonable Effectiveness of Mathematics in the Natural Sciences," the conviction has become an article of faith. It seems to me, though, that the "God is a mathematician" viewpoint is one of selective perception. The great swindle of introductory physics is that every problem has an exact answer. Not only that, students are expected to find it. Such an approach inculcates our charges with an expectation that is, in fact, exactly contrary to the true state of the world. Vanishingly few problems in physics have exact solutions and a physicist's career is one of finding approximations and hopefully not being too embarrassed by them.

In a freshman course we introduce the simple pendulum—nothing more

Tony Rothman lectures in physics at Princeton University. His most recent book is Sacred Mathematics: Japanese Temple Geometry *with Fukagawa Hidetoshi (Princeton University Press, 2008). Mail: Physics Department, Princeton University, Princeton, NJ 08544. E-mail: trothman@Princeton.edu*

The builders of the Tower of Babel, descendants of the Flood, were people of one language. When the people became isolated from each other by the confusion of tongues, further work became impossible and the people dispersed. When physicists detect that their models have become isolated from each other, with adjoining compartments that have no door between them, they pretend not to notice and keep on building. The Tower of Babel by Pieter Bruegel the Elder, 1563.

Erich Lessing/Art Resource, NY

than a mass on the end of a string that oscillates back and forth. Initially Newton's laws lead to an equation that is too hard to solve and so we admonish students to simplify it by assuming that the pendulum is executing small oscillations. Then the exercise becomes easy. Well, not only is the original problem too difficult for freshman, it has no exact solution, at least not in terms of "elementary functions" like sines and cosines. Advanced texts tell you that an exact solution does exist, but the use of the term *exact* for such animals is debatable. In any case, replace the string by a spring and the problem can easily be made impossible to everyone's satisfaction. One must distinguish the world from the description afforded by mathematics. As Einstein famously put it, *contra* Wigner, "As far as the laws of mathematics refer to reality, they are not certain; as far as they are certain, they do not refer to reality."

The maxim might be taken to heart a few weeks later in a freshman course when instructors introduce their students to Newton's law of gravity. The famous law works exquisitely well, of course, but a singular strangeness goes unremarked. According to the equation, if two objects become infinitely close to one another, the force of attraction between them becomes infinite. Infinite forces don't appear in Nature—at least we hope they don't—and we dismiss this pathology with the observation that real objects have a finite size and their centers never get so close to each other that we need to worry. But the first equation in any freshman electricity and magnetism course is "Coulomb's law," which governs the attraction or repulsion of electrical charges and is identical in form to Newton's law. Now, in modern physics we often tell students that electrons and protons are point particles. In that case, you really do need to worry about infinite forces and it is exactly this difficulty that led to modern field theories, such as quantum electrodynamics. Well Newton himself said, "*Hypotheses non fingo*": "Look guys, the equation works, usually."

Electricity and magnetism courses certainly hold their share of mysteries. The highlight of any beginning course on this subject, at least for the instructor, is Maxwell's equations, the equations that unified electricity and magnetism into electromagnetism. Soon after postulating them, we demonstrate to our students that light consists of traveling electric and magnetic fields, oscillating at right an-

gles to one another. We next assert that light exerts a pressure on matter; it is this radiation pressure that provides for the detonation of hydrogen bombs and the possibility of solar sails. A common explanation of light pressure in undergraduate texts is that the electric field of the light wave causes electrons to accelerate in one direction, then the magnetic field pushes them forward. Not only is this explanation completely wrong, despite its appearance in fifth editions of books, but to correct it requires introducing a famous ad hoc suggestion known as the Abraham-Lorentz model, which does reproduce the phenomenon. To put the Abraham-Lorentz model on a firm footing, on the other hand, led to the development of quantum electrodynamics. Quantum electrodynamics itself is, however, famously riddled with infinities, and to abolish them requires the further ad hoc procedure of renormalization, which was so distasteful to Paul Dirac that he ceased doing physics altogether. Although the theory of renormalization has advanced since those days, many physicists would echo Richard Feynman, one of the technique's inventors, who called it "hocus-pocus." Thus it is not entirely clear whether physics has ever provided an adequate underpinning to the wisdom so blithely dispensed in first-year texts.

Where the Action Is Not

One of the great moments in the lives of young physicists is their first encounter with "Lagrangian mechanics." God parts the firmament to reveal Truth. Lagrangian mechanics finds its roots in ancient Greece and in the 17th century suggestion of Pierre de Fermat that light travels between two points in the shortest possible time. Fermat's principle of least time allows one to derive the famous law of refraction, Snell's law and, apparently, explains the behavior of some dogs in retrieving bones. The idea that Nature minimizes certain quantities eventually led to the "principle of least action," which states that you take a quantity known as the action, minimize it and—*Presto!*—Newton's equations for the system miraculously emerge! (For classical systems the action is basically the system's kinetic minus potential energy, a quantity known as the Lagrangian, multiplied by the time.) The realization that Newton's laws themselves follow from the principle of least action is genuinely awe inspiring and the young physicist is immediately convinced that, logically if not historically, the action precedes Newton. Moreover, if you recollect (with pain) how difficult it was to even write down Newton's equations in freshman physics for those infuriating systems of ropes and pulleys, it becomes surprisingly easy in the Lagrangian formalism, as the technique is known.

But the great swindle of freshman physics is indeed the conceit that problems have exact solutions. A standard question during my sophomore year was to write down the equations for a double pendulum, which is merely one pendulum swinging from the tip of another. With the Lagrangian formalism, the task is a simple one. What went unsaid, or was perhaps unknown at the time, is that the double pendulum is a chaotic system, and so to solve the equations is strictly impossible. What then did we really learn from the exercise?

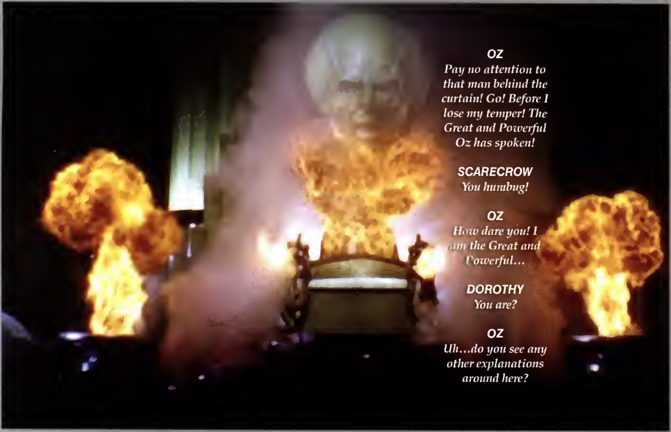
Nevertheless, in terms of its avowed purpose of deriving equations, the Lagrangian formalism works extraordinarily well for "standard systems" where the damned pulleys are connected by ideal, unstretchable ropes. Unfortunately, once the ropes are allowed to stretch with time, for example, the Lagrangian technique becomes anything but straightforward and the fixes needed to apply it to a given problem become so elaborate that virtually every textbook author either makes a misstatement or avoids those scenarios altogether. Indeed, the general situation is so delicate that one is often unsure whether one has obtained the correct result.

This is not a moot point. Einstein did not consider his theory of gravitation—general relativity—complete until he could derive his field equations from an action, a feat that the mathematician David Hilbert accomplished five days before Einstein himself. General relativity allows for many model cosmologies, most not resembling the real universe, but in any case today it is known that when deriving the field equations for certain of them from an action one gets an incorrect result. A fair amount of esoteric research has gone into understanding why the failure occurs and how to patch it up but, as far as I am aware, all fixes require assuming the correct answer to begin with—the Einstein equations. Modern physicists take the primacy of Lagrangian mechanics seriously: contemporary practitioners, be they cosmologists or string theorists, invariably begin by postulating an action for their pet theory, then derive the equations, but if one does not have a set of previously accepted field equations, how is one certain that one has obtained the correct answer, especially in this day when theory is so far removed from experiment?

It would be surprising if the strange world of subatomic and quantum physics did not lead the field in mysteries, conceptual ambiguities and paradoxes, and it does not disappoint. The standard model of particle physics, for instance (the one containing all the quarks and gluons), has no fewer than 19 adjustable parameters, about 60 years after Enrico Fermi exclaimed, "With four parameters I can fit an elephant!" Suffice to say, "beauty" is a term not frequently applied to the standard model.

One doesn't have to go so far in quantum theory to be confused. The concept of electron "spin" is basic to any quantum mechanics course, but what exactly is spinning is never made clear. Wolfgang Pauli, one of the concept's originators, initially rejected the idea because if the electron was to have a finite radius, as indicated by certain experiments, then the surface would be spinning faster than the speed of light. On the other hand, if one regards the electron as a point particle, as we often do, then it is truly challenging to conceive of a spinning top whose radius is zero, not to mention the aggravation of infinite forces.

Unfortunately, quantum texts habitually ignore the difficulties that infect the heart of the field. The most fundamental of these is the notorious "measurement problem." The equation that governs the behavior of any quantum system, Schrödinger's equation, is as deterministic as Newton's own, but as many people know, quantum mechanics predicts only the probability of an experiment's outcome. How a deterministic system, in which the result is preordained, abruptly becomes a probabilistic one at the instant of measurement, is the great unresolved mystery of quantum theory, and yet virtually none of the dozens of available quantum textbooks even mention it. One well-known graduate text completes



In science there is always another curtain to be pulled back.

Everett Collection

the irony by including a section titled "measurements" without addressing the issue.

Probable Cause

Quantum text authors, perhaps because of the perversity of their subject, are particularly adept at sweeping conceptual difficulties under the rug. Nowhere is this more apparent than in the celebrated "two-slit" experiment, which is universally invoked to illustrate the wave-particle duality of light and which brings you face to face with the bedrock inscrutability of Nature. The experiment is simple: Shine a light beam through a pair of narrow slits in a screen and observe the results. For our purposes, the great paradoxes illustrated by the two-slit experiment, that light can act like a wave or a particle but not both at the same time, are not central. What is central is that explanations of the experiment's results invoke both classical lights waves, on the one hand, and photons—quantum light particles—on the other.

Also central is that in analyzing this experiment textbook authors essentially throw up their hands and surrender. Recollecting that light is an electromagnetic wave, authors invariably begin by talking about the intensity of the incident light, which is a measure of the strength of the electric and magnetic fields. Then in a complete *non sequitur*, they shift the conversation to photons, as if the quantum-mechanical beastlets have electric and magnetic fields like classical light waves. They don't. In fact, an accurate description of the famous experiment requires a more subtle quantum-mechanical entity known as a coherent state, which is the closest thing to a classical light wave.

What's more, by resorting to a classical optics analogy of the experiment, authors are forgoing any explanation whatsoever. "Explanation" in physics generally means to find a causal mechanism for something to happen, a mechanism involving forces, but textbook optics affords no such explanation of slit experiments. Rather than describing how the light interacts with the slits, thus explaining why it behaves as it does, we merely demand that the light wave meet certain conditions at the slit edge and forget about the actual forces involved. The results agree well with observation, but the most widely used of such methods not only avoids the guts of the problem but is mathematically inconsistent. Not to mention that the measurement problem remains in full force.

Such examples abound throughout physics. Rather than pretending that they don't exist, physics educators would do well to acknowledge when they invoke the Wizard working the levers from behind the curtain. Even towards the end of the twentieth century, physics was regarded as received Truth, a revelation of the face of God. Some physicists may still believe that, but I prefer to think of physics as a collection of models, models that map the territory, but are never the territory itself. That may smack of defeatism to many, but ultimate answers are not to be grasped by mortals. Physicists have indeed gone further than other scientists in describing the natural world; they should not confuse description with understanding.

Bit Lit

Brian Hayes

Books are being blown to bits. New ones are "born digital"; millions of old ones are being assimilated into the mind of the machine.

Some people question the wisdom of this transition to digital reading matter. Paper and ink have served us pretty well for a thousand years or more. Is it prudent to store everything we know in tiny smudges of electric charge we can't see or touch? Critics also worry about who will wind up owning our cultural heritage. And then there are the sentimentalists, who say it's just not the same curling up by the fireside with a good Kindle.

Well, I for one welcome our new computer overlords. And I would like to point out that books are not *only* for reading. There are other things we (and our computers) can do with the words in books. We can count them, sort them, make comparisons among them, search for patterns in their distribution, classify them, catalog them, analyze them. Yes, these are nerdy, mechanical, reductionist assaults on literature—but they are also methods of extracting meaning from text, just as reading is. And they scale better.

Googling the Lexicon

The data-driven approach to language studies got a big boost last winter, when a team from Harvard and Google released a collection of digitized words and phrases culled from more than five million books published over the past 600 years. The text came out of the Google Books project, an industrial-scale scanning operation. Since 2004 Google Books has been digitizing the collections of more than 40 large libraries, as well as books supplied directly by publishers. At last report the Google scanning teams had paged their way through more than 15 million volumes. They estimate that another 115 million books remain to be done.

With digitized text from five million books, one is never at a loss for words

At the Google Books website, pages of scanned volumes are displayed as images, composed of pixels rather than letters and words. But to make the books searchable—which is, after all, Google's main line of business—it's necessary to extract the textual content as well. This is done by the process known as optical character recognition, or OCR—a computer's closest approximation to reading.

In 2007 Jean-Baptiste Michel and Erez Lieberman Aiden of Harvard recognized that the textual corpus derived from the Google Books OCR process might make a useful resource for scholarly research in history, linguistics and cultural studies. There are many other corpora for such purposes, including one based on a Google index of the World Wide Web. But the Google Books database would be special both because of its large size and because of its historical reach. The Web covers only 20 years, but the printed word takes us back to Gutenberg.

Michel and Aiden got in touch with Peter Norvig and Jon Orwant of Google and eventually arranged for access to the data. Because of copyright restrictions, it was not possible to release the full text of books or even substantial excerpts. Instead the text was chopped into "*n*-grams"—snippets of a few words each. A single word is a 1-gram, a two-word phrase is a 2-gram, and so on. The Harvard-Google database includes 1-, 2-, 3-, 4- and 5-grams. For each year in which an *n*-gram was observed, the database lists the number of books in which it was found, the number of pages within those books on which it appeared and the total number of recorded occurrences.

The *n*-gram database is drawn from a subset of the full Google Books corpus, consisting of 5,195,769 books, or roughly 4 percent of all the books ever printed. The selected books were those with the highest OCR quality and the most reliable metadata—the information *about* the book, including the date of publication.

A further winnowing step excluded any *n*-gram that did not appear at least 40 times in the selected books. This threshold, cutting off the extreme tail of the *n*-gram distribution, greatly reduced the bulk of the collection. Combining the 40-occurrence threshold with the 4 percent sampling of books, a rough rule of thumb says that an *n*-gram must appear in print about 1,000 times if it's to have a good chance of showing up in the database.

The final data set covers seven languages (Chinese, English, French, German, Hebrew, Russian and Spanish) and counts more than 500 billion occurrences of individual words. The chronological range is from 1520 to 2008 (although Michel and Aiden focus mainly on the interval 1800–2000, where the data are most abundant and consistent).

Culturomics

This past January, Michel, Aiden and a dozen co-authors from Harvard, Google and elsewhere published a research article in *Science* introducing the new linguistic corpus and presenting

Brian Hayes is senior writer for American Scientist. *Additional material related to the "Computing Science" column appears at http://bit-player.org. Address: 11 Chandler St. #2, Somerville, MA 02144. Internet: brian@bit-player.org*

some of their early findings. They also announced the Google Books Ngram Viewer, an online tool that allows anyone to query the database. Finally, they made the entire *n*-gram data set available for download under a Creative Commons license.

Some of the results reported in the *Science* article show how *n*-gram data can be used to document changes in the structure of language. One study examines the shifting balance between regular and irregular verbs in English—those that form the past tense with *-ed* and those that follow older or odder rules. Between 1800 and 2000 six verbs migrated from irregular to regular (*burn, chide, smell, spell, spill* and *thrive*) but two others went the opposite way (*light* and *wake*). In the case of *sneaked* vs. *snuck*, it's too soon to tell.

Michel and Aiden describe their work as *culturomics*, a word formed on the model of *genomics* (but not yet to be found in the *n*-gram data set). In the same way that large-scale collections of DNA sequences can reveal patterns in biology, high-volume linguistic data can aid the analysis of human culture. For example, Michel and Aiden examined changes in the trajectory of fame over the past two centuries by counting occurrences of celebrity names. According to the *n*-gram analysis, modern celebrities come to public attention at an earlier age, and their fame grows faster, but they fade faster, too. "In the future, everyone will be famous for 7.5 minutes," they remark (attributing the quote to "Whatshisname").

Another study looked at linguistic evidence of censorship and political repression. In English, the frequency of the name *Marc Chagall* grows steadily throughout the 20th century, but in German texts it disappears almost entirely during the Nazi years, when the artist's work was deemed "degenerate." Similar cases of suppression were found in China, Russia and the United States. (The American victims were the Hollywood 10—writers and directors blacklisted from 1947 until 1960 because of supposed Communist sympathies.)

Having found that known cases of censorship or suppression could be detected in the *n*-gram data, Michel and Aiden then asked whether new instances could be identified by searching among the millions of time series for those with a telltale pattern. In the case of the Nazi era, the team devised a "suppression index" that compares *n*-gram frequencies before, during and after the Hitler years. Starting with a list of 56,500 names of people, they found that almost 10 percent showed evidence of suppression in the German-language data, but not in English.

The Oracle of N-grams

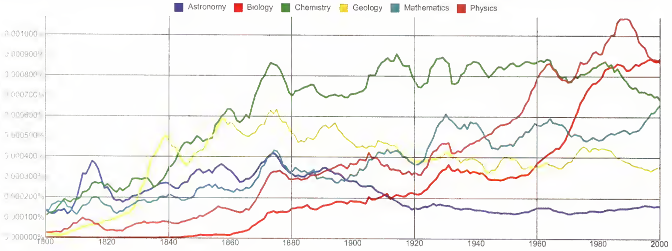
Reading about these experiments gave me the itch to try running some of my own. And it turns out that many interesting questions can be investigated with little effort and no cost using the Google Ngram Viewer (http://ngrams.googlelabs.com). The protocol for this service is simple: Type in a comma-separated series of *n*-grams, and get back a graph showing their frequency as a function of time. The frequencies are normalized to adjust for linguistic inflation—the expansion of the language as more books are published each year. The normalized frequency is the number of occurrences of an *n*-gram in a given year divided by the sum of all *n*-gram occurrences recorded in that year.

Shown below is the Ngram Viewer's output in response to a simple query—a list of six nouns. Interpreted with care, a chart like this one might tell us something about the shifting fortunes of scientific disciplines—but the careful interpretation is crucial. This is a popularity contest among words, not among the concepts they denote. From the graph it would appear that *Biology* did not exist before about 1840—and that's close to the truth if we're speaking of the word itself. But the science of living things goes back further.

The curves have some curious features that I can't explain, such as synchronized humps in about 1815 and 1875. Was there a real (but short-lived) upsurge in publishing books on the sciences in those years? Or are we seeing some artifact of librarianship or the selection process? The geology curve appears to have a persistent oscillation with a period of roughly 20 years. What, if anything, is that about?

The same query words without the initial capital letters yield somewhat different results. So do the corresponding agent nouns—*astronomer, biologist*, and so forth.

The Ngram Viewer can become an absorbing (and time-consuming) en-



Trends in word usage are traced across the centuries in a database built from the text of more than five million books, digitized as part of the Google Books project. The curves were drawn by the Ngram Viewer, an online tool available at http://ngrams.googlelabs.com. The frequency of each word is calculated as a proportion of all words published each year. For example, the word *Physics* was counted 55,757 times in books published in 1997; the total number of word occurrences in that year was 5,395,715,208, giving *Physics* a normalized frequency just over 0.001 percent.

tertainment. You might even turn it into a party game: One player draws the graph, the others try to guess the query. But less-frivolous applications are also within reach. Here's one possibility: With well-crafted queries, it might be possible to gauge the penetration of various foreign languages into English publications (or vice versa). From a very small sample, I get the impression that the frequency of German words in English text sagged during the World Wars, whereas Russian peaked in the Cold War.

Terror of Terabytes

The Viewer is an excellent oracle for the *n*-gram data, but it answers just one kind of question: How has the frequency of a specific *n*-gram varied over time? Many other questions cannot be expressed in this form. You might want to know which *n*-grams are the most common, or how word frequency varies as a function of word length, or which words entered the printed record first. To answer these questions and others like them, you'll have to work a little harder. For starters, you'll have to download the data, which is not a trivial undertaking.

The complete set of English *n*-grams weighs in at 340 gigabytes of compressed files, which expand to fill 2.5 terabytes of disk space. I have not yet tried to swallow all that data; doing anything interesting with it would require more hardware. I have been working solely with the English 1-gram files, which amount to 10 gigabytes when decompressed. I've been able to manage them on a laptop, although I've needed a refresher course in "external" algorithms—those that manipulate data on disk rather than in main memory. (This was a common practice when memory topped out at 64 kilobytes, but that was a long time ago.)

The 1-gram data are scattered over 10 files, which I merged into one. Then I set about gathering some basic facts and figures. In the English 1-gram data set there are 7,380,256 unique words, which occur a total of 359,675,008,445 times. Thus the mean number of occurrences per word is 48,735—but that's a somewhat misleading number, because the distribution is highly skewed. (The top 100 words account for half of all word occurrences.) A more meaningful statistic is the median, which is 166.

Which are the most common 1-grams? Setting aside a few common marks of punctuation, the highest-frequency words are: *the, of, and, to, in, a, is, that, for, was.* Another trivia question: What's the longest word in the corpus? I think the longest that's really a word and that wasn't invented just to set records is *phosphoribosylaminoimidazolecarboxamide.*

Prowling around in the data with a text editor reveals a multitude of oddities. Choose an entry at random, and it's likely to be a word you've never seen before. Indeed, there's a good chance it's not a word at all in the strict sense, but rather a number or a mixture of letters and digits, or something even more mysterious. For example, my eye fell on this curious "word":

BOBCATEWLLYUWXCARACALQW

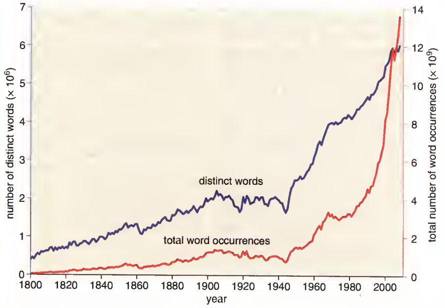
How could such a zany-looking string of letters turn up at least 40 times in published books? As it happens, we have a tool for answering such questions, namely Google Books. Since the Google OCR program produced this string, the Books search engine should be able to find it. And there it is: a row of letters in a word-search game—a game that has apparently been reprinted in dozens of puzzle books.

The Book of Numbers

On looking at the numbers included in the *n*-gram archive, I was surprised at first by their abundance. Of the 7.4 million unique 1-grams, about 7 percent are numbers or numberlike strings of digits. But the explanation is straightforward: Numbers have higher entropy than words. Only a tiny fraction of all possible sequences of letters make a meaningful word, but almost any combination of digits is a properly formed number. Thus for a given total quantity of numbers, we can expect to find greater variety.

To look more closely at the numeric 1-grams, I had to decide exactly what I would accept as a number. The OCR system allows mixed strings of letters and digits (1Deut, Na2SO4), but I wanted to consider just "pure" numbers, those that denote a definite numeric value or magnitude. I decided to accept any sequence of characters consisting entirely of digits or digits with a single embedded decimal point. The OCR program also accepts numbers preceded by a "$" sign, so I collected those dollar amounts too, but in a separate file.

Many different numerals can represent the same number: 01, 1, 1.0, 1.00 and 1.000 are all listed separately in the 1-gram files, but they all designate the same mathematical magnitude. I consolidated these items under the canonical value 1.0, and merged their



Overall growth of the English language can be seen in both the number of distinct words *(blue curve and left scale)* and the number of times those words appear in the corpus *(red curve and right scale)*. Both curves have a noticeable change of slope after World War II. The number of distinct words rises to 6 million in 2008, and the number of word occurrences approaches 14 billion. The database counts every spelling and capitalization variant separately, and also includes numbers, foreign words that appear in English books and much else.

yearly occurrence data into a single record. This procedure is not without drawbacks, in that it treats as numbers some items that aren't meant to designate a numeric value, such as Zip codes. But I don't know how to weed out those items.

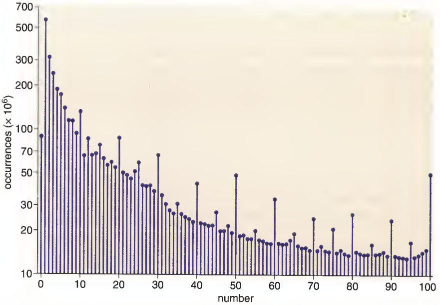
The number list I compiled has 458,794 unique values. The smallest is necessarily 0, since the OCR process strips away any minus signs. What's the largest entry? It's the number formed by repeating the digit 7 exactly 80 times. When I looked up the origin of this curious value, I discovered images of computer punch cards, with labeled rows of 80 columns.

The first thing I did with the numbers was check to see if they obey Benford's law, which describes the distribution of first digits in most of the numbers we meet in everyday life, such as those in stock-market tables. The law predicts that 1 is the most common leading digit, with higher digit values getting progressively rarer. In the theoretical distribution the frequency of digit d is proportional to $\log_{10}(1+1/d)$.

When I tested the 1-gram numbers against the predictions of Benford's law, the result was inconclusive. As expected, smaller first-digit values are more common among the 1-grams, but the preference for 1 is even more exaggerated than the Benford distribution predicts. The first digit should be a 1 about 30 percent of the time, but the actual frequency is 43 percent. Maybe those Zip codes are causing trouble?

I have another hypothesis: The distortion is caused by the times we live in! High on the list of popular numbers are values that look like years, almost all of which begin with 1. Numbers such as 2000, 1990, 1992 and 1980 are roughly 100 times more frequent than other four-digit numbers. To test my hypothesis I created an altered data set in which all numbers in the range 1800–1999 have their frequency artificially reduced by a factor of 1/100. The result is considerably closer to the Benford distribution, with 1 having a frequency of 34 percent *(see illustration at right)*.

Something else revealed by this collection of numeric data is the extraordinary human fondness for round numbers. The illustration at the top of this page plots the abundance of the first 100 integers. For the most part, frequency decreases with increasing magnitude, but numbers that are



Numbers included in the book-scan data reveal distinctive patterns of preference. Small numbers are more common than large ones, and round numbers (divisible by 10 or by 5) stand out above their neighbors. The data shown are for the integers from 0 through 100, but similar patterns are also observed elsewhere on the number line and at many scales.

"rounder"—divisible by 10, or if not by 10 then by 5—stand out above the crowd. (Also note that the integers 7 and 11, which by some vague measure might be taken as the *least* round numbers, are curiously depressed.)

Dollar amounts are even more dramatically biased in favor of well-rounded numbers. I had expected the monetary subset to be full of numbers ending in 99. Maybe that will be the case if we ever get an archive of junk mail and supermarket advertising, but in books there's a distinct preference for trailing zeros. The most popular dollar amounts are 1, 100, 2, 5, 10, 1000, 10000.

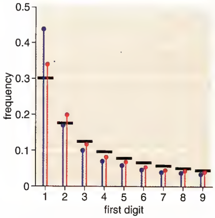
Quickfilver

Michel and Aiden set out to study language and culture, but they have also created a resource for the study of optical character recognition.

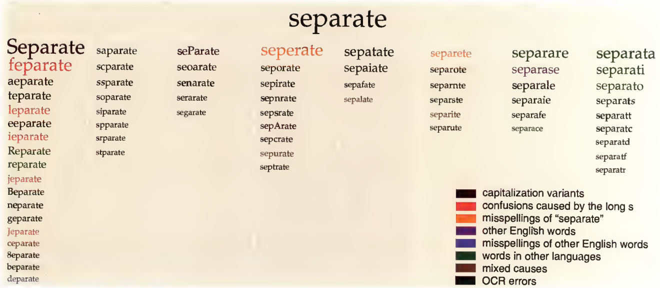
Based on a random sample from the 1-gram files, I estimate that 15 percent of the entries are affected in some way by OCR errors or anomalies. This sounds horrendous, but it does *not* mean that the OCR program made mistakes on 15 percent of the words it read; the word-recognition error rate is probably well under 1 percent. The problem is that there's only one way to read a word correctly, but there are countless ways to go wrong. Suppose the program reads 1,000 words and gets 990 of them right. If it makes a different mistake on each of the remaining 10 words, then the final list has 11 entries, 10 of which are erroneous.

Because of this effect, efforts to tidy up OCR errors would not only improve the accuracy of the data set but would also reduce its bulk. Entries for *Rccovery, Reeovery, Reoovery, Rerovery* and *Revovery* could all be merged into *Recovery*. But making such repairs is a daunting task, especially if you want to preserve other variations and errors, introduced not by the OCR process but by authors and printers.

Consider: *bomemaker* is probably an OCR error; *invertibrate* is probably



Leading digits of numbers in large collections are expected to follow Benford's law, defined here by the black stair steps. The Google Books numbers *(blue stalks)* have an excess of initial 1s. De-emphasizing year numbers of the 19th and 20th centuries yields a distribution closer to the expected shape *(red stalks)*.



A study of errors and oddities introduced by optical character recognition began by selecting a single word, *separate*, and then finding all other words in the database that can be produced by changing exactly one letter of *separate*; 65 such terms are present. Here the frequency of each word is encoded in its type size, which is proportional to the logarithm of the word's total abundance. (The main entry, *separate*, appears 27,528,661 times; the rarest forms, 42 times.) Consulting images of pages in Google Books showed which forms were likely to be OCR errors and which had other explanations. For example, *seperate* is a frequent spelling mistake, but *sepatate* and *scparate* are more likely to be machine misreadings. More than 60,000 errors are caused by mistaking the 18th-century long *s* for an *f* or another letter. And some of the variants are not errors at all: *separase* is the name of an enzyme, and *separare* is an Italian verb. Correcting all the errors would have only a tiny effect on the frequency ascribed to *separate*; on the other hand, it would purge the database of more than 40 spurious words.

a human error; *cerimoniale* is probably not English. What about *haccalaureate*? Is that an OCR error or is it a degree granted by a programming school? A human reader can make judgments in such cases, but hand-grooming multigigabyte files is not an attractive prospect. We need a mechanized solution.

A few special cases look doable. The OCR program has encoded some instances of "fi" and "fl" as ligatures, combining the two letters into a single character, while other instances remain as pairs of letters. For most uses of the data set, it would probably be better to treat all these cases consistently; this seems easy to accomplish.

More challenging but perhaps still within reach is the problem of "ſ," the "long s" that was part of English orthography through the 18th century. OCR programs (like many human readers) tend to interpret this character as the letter "f," leading to an abundance of comical fricative spellings such as *quickfilver* and *abfceffes*. I fufpect that an algorithm could fuccefffully correct a large fraction of these misreadings without turning too many *flaws* into *slaws*. The idea is to make a change only when the "s" form of a word is substantially more common than the "f" form and when the "f" version has a strong peak of popularity before 1800. But I have not yet tried implementing this algorithm, so I don't know how many new errors it will introduce.

With other OCR quirks, the chronological clue is lacking, and so we must resort to blunter tools such as a matrix estimating the probability that any one character will be mistaken for another. No doubt much can be accomplished in this way. On the other hand, if these mistakes were easy to fix, they would have been fixed already.

The Library of Babble

Suppose we had a magically cleansed version of the Harvard-Google database, free of all OCR errors. Would the time-series graphs from the Ngram Viewer look much different? I doubt it. Thus for present purposes the noise introduced by the OCR process is a minor distraction we can safely ignore.

But there are purposes beyond the present ones. Google has announced the grandiose goal of digitizing all the world's books. They may succeed. Some of those books may survive *only* in digital versions. And someone may even want to read them! If the scanning protocol now in use is the main channel by which we are to transmit 600 years of human culture to future generations, there's reason to worry.

But for the moment I am not inclined to complain. The *n*-gram collection released by the Harvard-Google team is a marvelous gift. I would much rather have it now than wait for some unattainable level of perfection. And now that it's been made public, it's ours as well as theirs, and we can all help improve it.

Bibliography

Abello, James, Panos M. Pardalos and Mauricio G. C. Resende (eds). 2002. *Handbook of Massive Data Sets*. Massive Computing, 4. Dordrecht: Kluwer Academic Publishers.

Michel, Jean-Baptiste, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, the Google Books Team, Joseph P. Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, Steven Pinker, Martin A. Nowak and Erez Lieberman Aiden. 2011. Quantitative analysis of culture using millions of digitized books. *Science* 331:176–182.

Norvig, Peter. 2009. Natural language corpus data. In *Beautiful Data*, edited by Toby Segaran and Jeff Hammerbacher, pp. 219–242. Sebastopol, Calif.: O'Reilly.

Smith, Ray. 2007. An overview of the Tesseract OCR engine. In *Proceedings of the Ninth International Conference on Document Analysis and Recognition, ICDAR 2007*, Vol. 2, pp. 629–633. New York: IEEE.

Vitter, Jeffrey Scott. 2008. *Algorithms and Data Structures for External Memory*. Series on Foundations and Trends in Theoretical Computer Science. Hanover, Mass.: Now Publishers.

Controlled Demolition

Henry Petroski

Removing structures can require nearly as much planning as building them

Most large structures, such as tall buildings, massive sports stadiums and long-span bridges, eventually reach or exceed their useful life and so need to be removed, either before or after a replacement is available. In skyscraper cities like New York, where land can be scarcer than clean air, it is not uncommon for a perfectly serviceable structure of modest height to be demolished to make way for a taller one. With sports stadiums in the age of television, it is not necessarily so much the need for a larger capacity as for more amenities, such as luxury skyboxes, that can motivate the destruction of a once-hallowed venue.

This was the case with the revered Yankee Stadium, "the house that Ruth built" in the 1920s. By the 1980s, there were calls for a new ballpark, with the Yankees' notorious owner George Steinbrenner claiming that the old one was unsafe. Indeed, in 1998, just days short of the stadium's 75th anniversary, a 500-pound piece of the structure fell without warning from beneath the third deck onto seats on the one below. Fortunately, there was no game being played at the time, so no one was hurt in the incident, but it did emphasize the deteriorating condition of the aging structure. After much posturing and political maneuvering to keep the team playing in New York City, construction of a new stadium began in 2006. The construction site was located on parkland just across the street from the old stadium, in which the Yankees continued to play ball. In 2009, with the completion of the new venue, which was designed deliberately to resemble the old, demolition of the obsolete one could commence.

Henry Petroski is Aleksandar S. Vesic Professor of Civil Engineering and a professor of history at Duke University. Address: Box 90287, Durham, NC 27708-0287

With the new stadium and the transportation infrastructure that essentially brought fans right up to the entrance of the old ballpark so close by, demolition had to proceed in a tightly controlled and low-risk manner, lest collateral damage be done. Heavy equipment staged on the famous ball field was used to tear down large sections of the structure at a time, pulling them inward and away from the new stadium and the elevated train station. Neighborhood residents, waiting for the site to be cleared so that they could regain a park, became impatient with the slow progress of the job, which was not completed until 2010. The use of explosives to implode the old stadium would certainly have speeded up the process, but at risk to the surrounding area. Fortunately, there are many cases where that risk is sufficiently low to allow the use of the more dramatic and expeditious means of demolition.

Deliberate Failure

Obsolescence and safety considerations can also demand the destruction of an old bridge that is too narrow or too weak for the expectations and demands of today's traffic. It is not that the bridge was improperly designed for its time; it is just that design criteria have evolved to meet new expectations or that the old design is obsolete in strength or size for increased traffic. But disassembling a sizeable steel bridge is no easy or risk-free task, and it can be a time-consuming and expensive undertaking. Thus, a common way to achieve the desired end—to remove the bridge in the quickest, safest and most economical way—is often to cause it deliberately to fail catastrophically and drop into the water below, from where if necessary the pieces can be dredged up and sold as scrap. A typical means for achieving such an end is to plant explosives at critical points on the structure and bring it down in a dramatic but controlled demolition.

This was the plan for removing two obsolete cantilever truss crossings—the 1929 John P. Grace Memorial Bridge and the 1966 Silas N. Pearman Bridge—with which the citizens of Charleston, South Carolina, had developed a love-hate relationship over the years. In 2005, after traffic was transferred to a new cable-stayed structure—the Arthur Ravenel Jr. Bridge—that had been built beside them, the main spans of the old Cooper River bridges were to be fitted with explosive charges that when detonated would cause failures to occur at critical locations and so bring the structures down. To establish the proper location, size and sequence of detonation of each of the charges to cause controlled failure was as much of an engineering problem as was designing and building the bridges to not collapse in the first place, and calculations and drawings signed and sealed by a professional engineer had to be prepared. The demolition process began with the removal of the signs, lights and other appurtenances on each bridge's roadway, followed by the removal of the roadway itself. Beams spanning over Drum Island, around which the flow of the river divides, were severed by machines with massive jaws and were allowed to drop onto the unoccupied island, from which they were removed. Concrete beams over populated areas and recyclable steel girders were care-

AP Photos/Bebeto Matthews

Most large structures have limited life spans and eventually must be removed. Such was case for the 1920s-vintage Yankee Stadium when a replacement was completed in 2009. The new ball park and an elevated train station were adjacent to the old stadium, however, so great care was required during demolition. Heavy equipment was used to pull the stadium inward, section by section, as shown here in March of 2010.

fully removed by cranes, which placed the salvaged parts on trucks. Where possible, high concrete piers were brought down by dynamite and then broken up on the ground. Explosives were also used to break up support structures located above and beneath the water.

To deconstruct the main spans of the bridges—the cantilever trusses over the navigation channels—their concrete decks were removed first, and then the stringers between the girders. This left a pair of lacey steel superstructures on which shaped charges were strategically placed so that the blasts would sever the skeleton spans into twenty- to forty-ton sections that could be retrieved from the water. To make that easier, the parts were rigged with retrieval cables and buoys before the explosives were set off. The first spans to be blasted into the water were those over Town Creek, which contained the side channel of the river. This provided an opportunity to test the scheme before repeating it over the river's main shipping channel. It was imperative that there were no surprises there, because the penalty for not opening that channel on time was $15,000 per hour or $360,000 per day late. It was fortunate that there was a practice try, because not everything went as planned: The designed failure scheme failed to go as planned. One large portion of the superstructure was left intact, and it took three weeks to recover all the steel out of Town Creek. An improved method on the creek's second bridge, which left it in larger pieces, enabled the retrieval process to go more quickly. The disappointing results with explosives led to an entirely different approach over the main channel of the Cooper River. Instead of blasting the bridge apart and letting the parts fall into the water, the bridge was cut up in place and the parts lowered onto barges. With parts as heavy as 600 tons, special heavy lifting and lowering equipment mounted on barges had to be used, resulting in what was a more expensive but also a more reliable operation. The steel trusswork was barged upriver to be cut apart and recycled. Concrete rubble was barged out to the ocean, where it was dumped in preapproved locations to create artificial reefs. Not all demolition jobs have such problems or complications.

The 80-year-old Crown Point Bridge was a vital link across Lake Champlain, connecting Fort Crown Point, New York, and Chimney Point, Vermont. As early as 2006, plans were announced to address the deteriorating condition of the bridge. In 2009, a routine inspection of the structure detected much larger cracks in its concrete piers than had been found in the prior biannual inspection, indicating a condition of accelerating deterioration subsequently confirmed by divers conducting an emergency underwater inspection. The concrete piers, which were designed and constructed to 1920s standards, were not reinforced with steel, as would be the practice today. Without reinforcing, the piers were less able to resist the push of ice, which tends to form first near the shore, where the water depth is less than what it is in the center of the lake, and so pushes unevenly on the piers. Such a scenario was evidently not anticipated by the original design engineers, and it was not made manifest before the unreinforced piers began to show signs of increasingly bad cracking.

Since the piers were so weakened by the cracks, the bridge itself was believed to be in danger of falling into the lake. Within weeks of the emergency inspection, a report was filed and the bridge was closed to traffic. Not only was the spontaneous breakup of its supporting piers a possibility, but also the aging steel superstructure had frozen bearings, a common ailment akin to the arthritic conditions that aging humans experience in their joints. The closing of the Crown Point Bridge creating considerable hardship for those who had relied upon it to commute to their jobs across the lake, for without the bridge they had to drive an extra hundred miles each way. To alleviate the inconvenience somewhat, an emergency project was undertaken to construct approach roads, parking lots, and other facilities to institute temporary ferry service until a new bridge could be completed.

A weak or faulty bridge naturally presents considerable risk to those who use it, and that risk can be completely eliminated by closing the bridge. But in the case of the bridge across Lake Champlain or over any recreational waterway, there is also the risk of a deteriorating structure falling on boaters. Since boat traffic moving under the thousands of feet of steel trusswork would be more difficult to control than the two lanes of road traffic that moved over and through it, the decision was made to demolish the bridge as soon as possible—preferably before the weather warmed up and the boats came out in force. In December 2009, the old bridge was brought down with the detonation of 500 charges packed with a total of 800 pounds of explosives. Such instant demolitions often make a big splash on the television evening news, and there is no shortage of YouTube videos of them on the Web. The debris that fell into Lake Champlain was to be removed by the spring, when construction of the new bridge was to commence, with a completion date set for 2011. The new bridge, which was variously described as a "modified network tied arch bridge" and a "steel structure with a handle-like arch along the main span," was designed to have a life expectancy of 75 years, which is remarkably close to the age to which the old span survived.

When Things Go Wrong

Sometimes, as with the Cooper River bridges, a planned failure of a land-based structure can be a failure itself. High old smokestacks, which may once have been landmarks, are often the object of demolition to clear the way for something new. Setting off a well-placed explosion or set of explosions at the base of the tall, slender structure is usually sufficient to topple it. When done correctly, the explosives will bring the stack down as effectively and precisely as a properly wielded power saw can a tree. When a chimney is made of bricks, knocking the base out from under it will result in a phenomenon that is often observed: while toppling, the chimney will break apart before it hits the ground. This happens because, while the brick structure may act like a monolith as long as it is standing up straight, when it is falling internal forces induced by the accelerating mass of the structure become greater than the mortar can bear. The chimney breaks where these forces first overcome the strength of the binding material. Exactly where the chimney breaks apart depends upon the shape and taper of the tower of bricks, and also upon how tenaciously the mortar holds the bricks together. For a typically shaped chimney with deteriorated mortar, the bricks will begin to separate very soon after the chimney begins to topple and close to halfway up; for a chimney in which the mortar is strong, the break will more likely occur later in the fall and closer to about one-third of the way up from the bottom.

Some tall chimneys are made of reinforced concrete, in which case the acceleration forces generally cannot become large enough to cause the concrete to separate during the structure's fall. This is exactly why reinforcing steel is embedded in the concrete in the first place; it helps keep them from cracking in the wind. When such a chimney is demolished, it will typically fall in one piece, perhaps breaking up somewhat only when it strikes the ground. It is very important, of course, for the demolition explosives to act on the base in just the right way so that the chimney falls in the right direction. This did not happen at the 90-year-old Mad River power plant in Springfield, Ohio, where in 2010 a 275-foot-tall concrete chimney was being demolished. Instead of falling to the east, as planned, the chimney fell to the southeast, severing two high-voltage power lines and damaging a pair of turbines at the plant, causing a power outage for about 4,000 customers. Speculation was that there were preexisting but undetected cracks in the chimney, which were not taken into account when placing the 17 pounds of explosives that were used. According to the "chief blaster," it was only the fourth time in his 31 years of imploding structures that the job did not go as planned. In fact, it was only the second

AP Photos/Mike Groll

The Lake Champlain Bridge, which connected Crown Point, New York, and Chimney Point, Vermont, was closed in October of 2009 because of severe cracking in its concrete piers. On December 28, 2009, the bridge was dropped into the lake by 500 charges totaling 800 pounds of explosives. The debris was to be removed by spring, when boat traffic would return to the lake.

AP Photos/Springfield News-Sun, Bill Lackey

Occasionally demolitions don't go quite as planned. When crews toppled the 275-foot smoke stack at Ohio Edison's Mad River power plant in Springfield, Ohio, on November 10, 2010, it fell to the southeast, rather than the east, taking out two high-voltage power lines serving some 4,000 customers.

time this had happened in his last 25 years of work. As evidence of his expertise, he explained, "I just successfully 'shot' an 85-foot-tall building in Athens, Greece, with only 15 feet of clearance on two sides and 30 feet of clearance on the other two." That's the way demolition is supposed to work—and usually does.

Old or defective buildings of all kinds are frequently the object of designed failures. The 376-foot-tall Ocean Tower high-rise condominium on South Padre Island, Texas, was never occupied, because the structure began to lean even before it was completed. Uneven ground settlement was blamed for the problems that the uncommonly tall structure experienced on the sandy beachfront location, and the owner of the building decided to bring it down. The company Controlled Demolition was contracted to implode the building without damaging nearby houses, sand dunes and a park. The demolition firm described the task, which would set a height record for an imploded reinforced-concrete building, as "one of its most challenging razings using explosives." In fact, it considered the job to be "its second most difficult implosion after Seattle's Kingdome," the stadium that in 2000 was brought down by over two tons of dynamite.

The Texas island job was complicated for a variety of reasons. First, there was the height and slenderness of the structure, which meant that there was not much opportunity to blast its sides inward before the tower fell in the clutches of gravity. Further complications revolved around the fact that the three-sided condominium tower was built on top of a rectangular multistory parking garage, whose structural beam system included a weak (open) section that had allowed for passage through the floors of the base of the tower crane used during construction. Additionally, the tower was founded on sand, which is a good conductor of vibrations that can do damage to nearby structures. In order to deal with all of these complications and at the same time ensure that the falling building and debris would be contained within a limited area and would not rain down on neighboring property or onlookers, careful design and planning had to take place. The 2009 demolition of Ocean Tower proved to be a model of thoughtful design, planning and execution of a controlled failure. Strategically placed charges, backed up by debris-containing screens made of geotextiles—mat-like materials used to confine soil and rock—were set off in a precisely timed sequence that in the span of 12.5 seconds straightened the tower, leaned it about 15 degrees toward the Gulf of Mexico, and then let it drop vertically into a well-contained pile of broken concrete and twisted steel. It was indeed a successful failure.

As the 1993 explosion in the basement public parking garage of the north tower of New York's World Trade Center demonstrated, even a truck full of explosives may not be enough to cause a building to fail. The truck bomb did do a lot of damage, taking out multiple basement floors and leaving a gaping hole behind, but the building itself did not come down the way the terrorists evidently had hoped it would. To bring down a building in a controlled (or even uncontrolled) failure mode takes a lot more than just a truck full of dynamite or other explosive. It takes strategically placed charges that are detonated in the proper sequence. In other words, to cause a structure to fail in a certain way is as much of a design problem as it was to ensure that the building stand up properly in the first place.

Conspiracy Against Oneself

In 2001, the collapse of the New York World Trade Center towers after being struck by hijacked airliners looked to some observers to be very much like controlled demolition. Indeed, conspiracy theorists attributed the cause of the serial collapses to the detonation of surreptitiously but well-placed explosive charges, offering as supporting evidence what were described as puffs of smoke that were emitted during the pancaking of the towers' floors. The collapses were certainly reminiscent of controlled building demolitions, but that does not mean they were. On the contrary, once fire had sufficiently softened what steel columns remained after the planes flew into the towers, the way they collapsed progressively under the weight of their own falling floors was evidence of why building demolitions work so successfully. A falling building effectively demolishes itself, and so the design problem faced by a "failure engineer"—if that term may be allowed—is to initiate the collapse in the most efficient and effective way and then let gravity do the heavy work.

The design of a benevolent demolition involves more than just establishing the size and location and firing sequence of explosive charges. A lot of the design involves preliminary work in removing or weakening strategic columns and other structural impediments to progressive collapse. Some columns are cut partially or completely through to insure that they will be moved sideways (and so out of the way of falling floors) by the action of the explosions. Support walls are removed so that they will not offer any resistance to the collapse. The building, in other words, is reduced to a weakened structural frame before any

Jess Merrill/Alamy

The Ocean Tower high-rise condominium building at South Padre Island, Texas, was never finished because uneven settlement caused it to lean. Strategically placed charges, detonated sequentially over a 12.5-second time span, brought the 376-foot-tall building down upon itself. The technique, referred to as implosion, employs the building's own mass to hasten its demise.

explosives are even put in place. In the case of the September 11 attacks, the fires that were ignited by burning jet fuel and fed by office furniture and supplies did the work of weakening the structural columns that had not been taken out by the plane impacts. The temperature of the flames naturally heated up the steel columns, which in time became softened to the point where they could not support the load that bore down on them. As a column gave out, its weakened neighbors were immediately required to take on additional loads, but because they too had been heated and softened they could not do so. Once the columns on the impact-damaged floors had reached the limits of their strength to hold up the floors above, those floors fell with a force greater than the columns on the floor below could bear, so the buildings collapsed progressively down to the ground. As floor fell onto floor, considerable impact forces crushed anything between them, and dust and fine debris were pushed outward by the rapidly compressed air. No explosives were needed, because the heat of the fires weakened critical support columns and the momentum of the falling floors provided sufficient energy to pulverize concrete and push steel columns aside.

Whether the terrorists who planned and executed the 2001 attacks on the World Trade Center had designed them as carefully as a professional demolition team would have to bring down such buildings may never be known for sure. What does seem likely, however, is that the terrorists did learn from prior failures. The failure of the 1993 basement truck-bomb explosion to cause the north tower to topple into the south one taught the evildoers that another strategy was in order. The terrorists may well have designed their airplane attacks to bring down the buildings, but they might not have known exactly where to impact them for the maximum likelihood of success. Seeing that the crash of the first plane did not cause the building hit to fail catastrophically, the hijacker piloting the second may have reasoned that he should direct its impact at a lower floor, so that there would be a greater portion of the building's weight bearing down on the severed columns, which numbered many more than the truck bomb had damaged. That this was ultimately a diabolically effective decision was demonstrated by the fact that the tower struck second was the first to fall.

In some cases, things can have been designed so conscientiously against failure that they present challenges when they have outlived their usefulness. Nuclear power plants, because of their potential to wreak so much havoc should an accidental release of radiation occur, have been built with extreme strength—enough in some cases to withstand the impact even of a large airplane. This resistance to failure, coupled with the residual radioactivity in decommissioned plants, has presented engineers and society with a new challenge: how to disassemble a plant that is no longer needed or wanted. But for all their thick reinforced-concrete walls, nuclear containment structures also have their failure points that will yield to properly designed and chosen wrecking tools, albeit ones that may have to be wielded by robots. One method that is generally not appropriate for an old nuclear power plant is demolition by explosives, however, for no matter how well-controlled, the explosive nature of the method is too likely to spread radioactivity in an uncontrolled manner.

Bibliography

Cho, Aileen. 2009. Officials hurrying with plans to replace closed crossing, *Engineering News-Record*, November 20, p. 15.

Dwyer, Charles. 2007. Cooper River bridge demolition, 36th Annual South Carolina State Highway Conference, Clemson, S.C., annotated slide presentation, http://www.clemson.edu/t3s/workshop/2007/Dwyer.pdf.

Larson, Mike. 2010. Tower knockdown scheme does not go as planned, *Engineering News-Record*, Nov. 22, p. 14.

McKinstry, Lohr. 2010. Flatiron wins contract for new Champlain bridge, *ENR.com*, May 29, http://enr.construction.com/yb/enr/article.aspx?story_id=145562812&elq=182e4fdeabe54cb69a1945be86b3afc7.

New York State Department of Transportation. 2010. Lake Champlain bridge project, https://www.nysdot.gov/lakechamplainbridge/history.

Post, Nadine M. 2009. Faulty tower's implosion will set new record, *Engineering News-Record*, Nov. 30, pp. 12-13.

Simanek, Donald. 2007. Physics lecture demonstrations, with some problems and puzzles, too, http://www.lhup.edu/~dsimanek/scenario/demos.htm, s.v. "Chimney toppling."

South Carolina Department of Transportation. 2004. Cooper River Bridge: History, http://www.cooperriverbridge.org/history.shtml.

Varieschi, Gabriele, Kaoru Kamiya and Isabel Jully. Accessed Dec. 2010. The falling chimney web page, http://myweb.lmu.edu/gvarieschi/chimney/chimney.html.

WPTZ.com. 2009. Controlled explosives topple aging Champlain Bridge, WPTZ.com, Dec. 21, http://www.wptz.com/print/22026547/detail.html.

Jefferson's Old Bones

Keith Thomson

IN APRIL 1836, the naturalist John Kirk Townsend wrote from Vancouver to the physician Samuel G. Morton in Philadelphia, to report that a group of Indians had told him of seeing

> a quadruped of a gigantic size which from the description appeared to resemble the supposed extinct *Mastodon.* They said it was about the height of one of their houses (not less than 30 feet), and that it was of a brownish or blackish color, & appeared almost destitute of hair.... This story was substantiated by the whole party, & some individually did not hesitate to strengthen their affirmations by the most solemn oaths.

Similar tales of strange monsters roaming the western lands of the continent had been common in the previous century. A nice example is given in the journal of James Kenny, a frontier trader who worked for the Commissioners for Indian Affairs, from 1761:

> the Rhinosses or Elephant Master, being a very large Creature of a Dark Colour having a long Strong horn growing upon his nose (wth which he kills Elephants) a Short tail like an Elk; two of sd horns he seen fixd over a Gate at St Augustine, & that its ye bnes of Some of these lies down in Buffelo lick by ye Ohio, wher ye Great teeth Comes from.

Thomas Jefferson avidly collected such accounts as they were important

Did the so-called father of American vertebrate paleontology believe in fossils?

for his view of science. Jefferson did not believe in extinction. He was particularly fascinated by the American mastodon, the elephant relative that he referred to for many years as "the mammoth." It was not until 1806 in Paris that the French naturalist Georges Cuvier formally separated "mastodonte" from mammoth and also concluded that there were two living species of elephant. But in his *Notes on the State of Virginia* (1785), Jefferson had already concluded that the cold-adapted "mammoths" were different from the living tropical African and Asian elephants. Over many years he amassed a large collection of "mammoth" remains, which he displayed in the entrance hall of Monticello, his great house in Virginia.

Jefferson's second great paleontological interest arose in 1799 when he was sent some bones of a huge mammal "of the clawed kind" that had been found in a cave in Greenbrier County (in present-day West Virginia). He described this creature in his writings and named it *Megalonyx* ("great claw"). Jefferson was sure that this creature also still lived and, as is well known, he very much hoped that the Lewis and Clark Expedition of 1804–1806 would find evidence of both living mastodons and *Megalonyx* in the American West. He was therefore pleased when, soon after he completed his study of *Megalonyx* for publication in 1799 in the *Transactions of the American Philosophical Society*, a correspondent informed him:

> Some circumstances have lately been related to me which in my opinion go far in support of your conjecture that the species of Animal whose Bones were found in Green, still exists in the Western Country.... far up the Misouri river in which an Animal is found of a brown colour, much larger than a Bear, of astonishing strength, activity, & fierceness ... a Nail [was] taken from the Claw of one killed by the party of Indians to which he belonged but not before it had torn several of them into pieces. This Nail or horny part of the Claw is said to measure six inches in length.

Father or Not?

For his analysis of the mastodon and description of *Megalonyx*, historians have often called Jefferson the father of American vertebrate paleontology. Sixty-five years ago, however, paleontologist George Gaylord Simpson, in a masterly review of the history of American vertebrate paleontology, argued that Jefferson did not deserve this accolade because his actions were not sufficiently scientific. Simpson points out that Jefferson did not believe in extinction partly on religious grounds, and in his paper on *Megalonyx*, he started with a theory (that the animal was some kind of gigantic American lion) and then tried to prove it, rather than first assembling the facts. Both of these charges are true. But the situation turns out to be rather more complicated than Simpson thought.

The first defense against Simpson's charge is that Jefferson's views on extinction were not simply, and certainly not only, a matter of religious prejudice—although it is true that Jefferson was a deist who thought that the

Keith Thomson is professor emeritus of natural history at the University of Oxford and senior research fellow of the American Philosophical Society. His recent books include The Young Charles Darwin *(Yale University Press, 2009) and* A Passion for Nature: Thomas Jefferson and Natural History *(Thomas Jefferson Memorial Foundation and University of North Carolina Press, 2008). Address: American Philosophical Society, 105 South Fifth Street, Philadelphia, PA 19106. E-mail: kthomson@amphilsoc.org*

Thomas Jefferson avidly collected evidence, in the form of both first-hand accounts and bones, of mastodons and *Megalonyx*, a prehistoric giant ground sloth (shadowing Jefferson in this artist's rendition). His insightful written analyses of these remains has led historians to call Jefferson the father of American vertebrate paleontology. But Jefferson did not refer to the bones as fossils, and he did not believe in extinction. He instead hoped to find proof that these creatures still existed in the American West. For these and other reasons, some noted paleontologists have argued that Jefferson did not deserve the honorific title.

Creator had made the world just once, and as it is now. He wrote to John Adams in 1823 of his beliefs:

> The movements of the heavenly bodies, so exactly held in their course by the balance of centrifugal and centripetal forces, the structure of our earth itself, with its distribution of lands, waters and atmosphere, animal and vegetable bodies, examined in all their minutest particles, insects mere atoms of life, yet as perfectly organized as man or mammoth, the mineral substances, their generation and uses, it is impossible, I say, for the human mind not to believe that there is, in all this, design, cause and effect, up to an ultimate cause, a fabricator of all things from matter and motion, their preserver and regulator …

In fact, Jefferson did recognize the obvious fact that species and populations—such as the wolf and bear in Britain, or various American Indian groups—became extinct. He also believed that nature compensated for such losses. The sentence quoted above continues, "while permitted to exist in their present forms, and their regenerator into new and other forms." Furthermore, in the same letter he wrote:

> We see … evident proofs of the necessity of a superintending power to maintain the Universe in its course and order … Certain races of animals are become extinct; and, were there no restoring power, all existence might extinguish successively, one by one, until all should be reduced to a shapeless chaos.

As for the mastodon and *Megalonyx*, Jefferson, the lawyer, simply argued "the bones exist; therefore the animal has existed.… If this animal then has once existed, it is probable … that he still exists." But he also argued like a scientist. In his paper on *Megalonyx* he devoted a full four (out of 14) pages to reports from western travelers of encounters like those cited above. In this sense, his view of extinction can be seen as a hypothesis for which he cited evidence in support.

A more difficult question concerns Jefferson's understanding of what his mastodon and *Megalonyx* bones actually were. It is impossible to open any work about Jefferson and his collections without finding a reference to Jefferson's love of "fossils." But a careful search of Jefferson's writings (now made possible by the availability of searchable databases) reveals the

Ted Daeschler / Academy of Natural Sciences

Jefferson depicted nine bones of the left paw of *Megalonyx* in his 1799 paper, photographs of which are show at left. An illustration of the *Megalonyx* paw indicates the placement of the bones *(right)*. The clawed digits were probably adapted for hanging from trees, and possibly for pulling down branches and stripping leaves. The two outer digits (IV and V) helped to bear the animal's massive weight, as they would be curled under the rest of the paw when it walked on the ground. (Illustration adapted from a drawing by Dennis Murphy/Academy of Natural Sciences.)

surprising fact that he never referred to them as fossils. For him they were always simply "bones." The word "fossil" does not appear in *Notes on the State of Virginia* or his letters.

Jefferson and "Fossils"

In Jefferson's day, the word "fossil" was undergoing a change of meaning. Prior to the mid-18th century, "fossil" (from its Latin root, *fossere*, to dig) referred to anything dug up out of the ground. Thus coal, diamonds, metal ores, asphalt and the remains of once-living creatures were all called fossils. Then in works such as John Woodward's *Fossils of all Kinds, Digested into a Method Suitable to their Mutual Relation and Affinity* (1728) two different categories came to be distinguished: *native fossils* (minerals and metals) and *extraneous fossils*. The latter was restricted to the remains of once-living creatures, petrified and preserved in the rocks. Eventually the various elements of native fossils all found their own separate names and the term *extraneous fossils* (and sometimes *adventitious fossils*), was reduced simply to "fossils," remained for the organic remains. From then, it was used both as a noun and an adjective. Thus in 1758, physician John Fothergill reported on "the fossile bones of an Allegator" from Whitby in Yorkshire, England, and anatomist William Hunter, writing about remains from Big Bone Lick in America in 1768, called them "fossil tusks." Botanist Peter Collinson the same year referred to "very large Fossil Teeth" from the same source.

It is important to note, however, that at first the definition of something as a fossil did not necessarily imply great age; the whole Earth, in conventional thinking, was only 6,000 years old. But, in that context, something 5,000 years old was ancient indeed. This still left the problem of discovering how petrification occurred and how fossils were formed—how "the stoney Atoms have intruded themselves into all parts alike,"

Ted Daeschler / Academy of Natural Sciences

Jefferson's collection of bones included this mastodon molar tooth from an older adult animal. Jefferson referred to this animal for many years as a mammoth, but in a 1785 publication, he concluded that the cold-adapted mammoths were a different species than living tropical Asian and African elephants. It was not until 1806 that a French naturalist formally separated mastodons from mammoths and concluded that living elephants were divided into two species.

as Robert Plot put it in his 1677 publication *Natural History of Oxford-shire.*

In America the transition to the new set of meanings was patchy. In 1771 the American Philosophical Society proposed the creation of a new museum containing "all Specimens of Natural Productions, whether of the ANIMAL, VEGETABLE or FOSSIL Kingdoms." A decade later, philosopher Pierre Eugene du Simitière advertised his "American Museum" in Philadelphia as including "Artificial Curiosities" (such as portraits and machines) and "Natural Curiosities." The latter included "Marine Productions" (fishes and the like), "Land Productions" (birds and insects, for example), "Fossils" (used in the old sense of minerals and salts) and "Petrifications." This last category included both what we now call fossils (such as bones, shells, teeth and corals) as well as a number of inorganic categories such as "fossil substances produced by the eruptions of Volcanos." As late as 1792, when painter and naturalist Charles Willson Peale proposed his new museum for Philadelphia, he wrote that it would include "the fossil kingdom, comprehending the earths, minerals, and other fossil matters, which include petrefections." All three projects set a line between objects that belonged to living organisms and those that were petrified.

An obvious logical difficulty created by such classifications was that an object such as a mastodon femur could be classed either as a bone or a fossil, but it could not be both. In this situation, anything that recognizably belonged to a vertebrate skeleton, even if petrified, was preferentially referred to as a "bone" or "tooth."

A modern usage of the word fossil appeared in the minutes of the American Philosophical Society in 1784, when a collection of Big Bone Lick fossils collected by a Major Craig was described as "petrified bones." The 1799 volume of the *American Philosophical Society Transactions,* which included Jefferson's *Megalonyx* paper, also had a contribution from a land speculator, Judge George Turner, in which bones from Big Bone Lick were termed "extraneous fossils." In a 1796 letter to his cousin Philip Turpin—a farmer whose land became an archeological site—Jefferson himself referred to an unspecified "petrified Bone in my possession" (evidently not part of his *Megalonyx* suite). Several of Jefferson's correspondents in the 1780s and 1790s used the word fossil in its current sense. He was also familiar with the fossil shells (possibly Silurian brachiopods) that could be found in the Blue Ridge Mountains of Virginia and freely referred to them as petrified but, again, not as fossils.

Jefferson's consistent failure to use the word in his more formal writings almost seems like deliberate avoidance. That impression is strengthened by the only two cases that I have discovered of Jefferson actually mentioning the word. In 1807 he hired William Clark to head back to the Ohio territory and make a large collection of bones from Big Bone Lick, the famous site from which all the great mastodon remains had been recovered (before Peale's then-recent excavations in New York State). When first formally commissioning Clark, Jefferson referred only to collecting "bones." When he thanked him for the results two years later, he also wrote of "bones." But in his private account book he made a different annotation. The entry for February 9, 1808, reads: "Gave ord. on bank US. In favr. Genl. Wm. Clarke for expenses digging fossil bones." There is a similar notation for November 2 of that year where he used the French spelling "fossile."

At the very least these entries tell us that Jefferson recognized the status of some remains as having been "dug up," but from his other writings it is not at all clear that Jefferson thought that his mastodon and *Megalonyx* remains were actually petrified. He wrote that the *Megalonyx* remains showed "a small degree of petrification" in addition to being preserved by impregnation by "nitre" from the cave floor. He may even have doubted whether elements found in a marsh or on a cave floor qualified as having been "dug up."

Scientist and Lawyer

It is inappropriate to judge Jefferson, as Simpson did, by our own standards, and many a great scientist turns out to have been (in our estimation) wrong some of the time. Nonetheless, as is often the case with Jefferson, a puzzle remains. Did he deliberately avoid using the word "fossil" in reference to the mastodon and *Megalonyx*? If so, did he do it as a lawyer and for strategic, rhetorical reasons, or as a scientist, or as both? From our modern point of view, the dilemma Jefferson faced is easy to outline. There was no consensus on how petrifaction occurred, how long it took, and indeed, how sea shells came to be found high up in mountains, how mountains were formed or how old the Earth was. And there was no unanimity about whether extinction was a real phenomenon. In each of these questions, some contemporary philosophers were edging toward modern solutions, whereas others were exploring versions of a flood theory or hypothesizing about impacts with comets. Jefferson, aware of the trends, evidently found his ideas and information to be in conflict.

From a religious point of view, Jefferson did not want to accept extinction, but there was also enough anecdotal evidence to suggest that *Megalonyx*, at least, might still be alive in the west to support him in that conclusion. He was unsure to what extent his "bones" were petrified. It seems that his solution was to be extremely careful in his word usage. As a master of language and of precise use of rhetoric, Jefferson seems to have kept his options open by simply avoiding use of the word fossil and by sticking to terms whose meaning was precise and unequivocal. By referring to the fossils of mastodon and *Megalonyx* (which in any case, we now know were only 8,000 to 12,000 years old) simply as bones, tusks and teeth, he wrote with strict accuracy and preserved his philosophical position.

With his perceptive analysis of the mastodon and other elephants in *Notes on the State of Virginia* and his account of *Megalonyx*, Jefferson made significant contributions to scientific knowledge. He certainly encouraged the discovery and study of fossils by people such as Charles Willson Peale and physician Caspar Wistar (whose own analysis of *Megalonyx*, published alongside Jefferson's, was a masterpiece of forensic anatomy). Jefferson was misled by his hopes that the mastodon and *Megalonyx* were mighty ferocious carnivores that would symbolize American vigor (and contradict European claims of the inferiority of American wildlife), but he was in many ways the progenitor of our modern fascination with another set of monsters from the west—dinosaurs. In these respects, he can fairly be considered a founder of American vertebrate paleontology. But the irony remains that he helped found a discipline without having accepted two of its fundamental premises: fossils and extinction.

Bibliography

Simpson, George Gaylord. 1942. The beginnings of vertebrate paleontology in North America. *Proceedings of the American Philosophical Society* 86:153–188.

Authorship Diplomacy

Melissa S. Anderson, Felly Chiteng Kot, Marta A. Shaw,
Christine C. Lepkowski and Raymond G. De Vries

Cross-national differences complicate allocation of credit and responsibility

Among scientists, authorship is a very big deal—and for good reason. It not only establishes the record of scientific progress but also stakes a scientist's claim to originality and priority. As sociologist Robert Merton noted decades ago, recognition for original work is the coin of the realm in science. Authorship is the basis for promotion, tenure, salary, honors and invitations to participate in prestigious initiatives. It is important for collaborating authors to get it right.

Getting it right seems like a simple and straightforward task: Include those who contributed to the project and omit those who did not. Most scientists, however, have encountered situations in which coauthors disagreed about who should be included on a publication or in what order they should be listed. In a recent study, two of us and our colleague Brian Martinson found that 12 percent of midcareer scientists admitted that they had inappropriately assigned authorship credit within the previous three years. Such situations may reflect competitive pressures in science or disputes among authors.

Melissa S. Anderson is a professor of higher education in the Department of Organizational Leadership, Policy and Development at the University of Minnesota, Minneapolis. She received a Ph.D. in higher education from the University of Minnesota. She studies research integrity and international scientific collaboration. Felly Chiteng Kot is a Ph.D. candidate and Marta A. Shaw and Christine C. Lepkowski are Ph.D. students in Anderson's department. Raymond G. De Vries is a professor of bioethics at the Center for Bioethics and Social Sciences in Medicine at the University of Michigan. He received a Ph.D. in sociology from the University of California, Davis. He is writing a social history of bioethics and researching the internationalization of ethics guidelines. Address: Melissa S. Anderson, Department of Organizational Leadership, Policy and Development, University of Minnesota, 330 Wulling Hall, 86 Pleasant Street SE, Minneapolis, MN 55455-0221. E-mail: mand@umn.edu

Problems with authorship are complicated enough in domestic research, but they can be particularly thorny in the context of international scientific collaborations. Whether authorship disagreements are more common in international or domestic research is an open question, but some aspects of cross-national collaboration do complicate authorship decisions.

International Collaborations

Scientific research is increasingly international in scope and practice. Worldwide, the percentage of science and engineering research articles with authors from more than one country increased from 8 percent in 1988 to 22 percent in 2007, according to the 2010 *Science and Engineering Indicators* compiled by the U.S. National Science Foundation. Rates of international collaboration as defined in the *Indicators* are 20 to 30 percent in the United States, China, Japan and India, but around 50 percent in the European Union, in part because recent EU policies and incentives favor international collaboration.

As we consider authorship issues that arise in these collaborative ventures, we draw on our own and our colleagues' work in the recently published book *International Research Collaborations: Much to be Gained, Many Ways to Get in Trouble* and our ongoing research on international scientific collaborations. Specifically, we use material from 10 focus groups and 60 interviews that we conducted over the past year with scientists in the U.S. (and a few outside the U.S.) who are involved in cross-national research collaborations. When we asked these scientists about problematic and beneficial aspects of international research, we inevitably heard about issues with authorship and publication.

Errors of Omission

One of the most obvious problems in collaborative authorship is omitting authors from a paper. The classic form of omission occurs when two collaborators are in conflict (professional or personal) and one leaves the other's name off a paper out of spite. Such cases are possible in almost any collaboration—domestic or international. But other forms of omission are more directly linked to cross-national research. Qualifications for authorship, based on scientific contribution or professional status, differ internationally. One scientist told us about working with collaborators in another country who were unwilling to give authorship credit to graduate students simply because of their junior status. He tried to correct the injustice without triggering professional retaliation against the students: "You really almost have to be subversive to help younger people in a way that doesn't ruin their lives at home, which is not so simple—but I think it is a huge integrity issue."

We also heard about researchers who left others' names off publications in order to advance their own careers. In some countries, a senior scientist may feel entitled to take full credit for a junior colleague's work. But the reverse can also happen: Sometimes young scientists train in labs outside their native countries, then publish the results on their own once they return home.

Authors who are omitted without having given their consent often feel wronged, but sometimes authors agree to be left off a publication in exchange

In this map of research collaborations, coauthors' cities are connected by white lines. The brightness reflects both the number of collaborations between a pair of cities and their proximity. The network is based on author affiliations on scientific papers aggregated in Elsevier's Scopus database from 2005 to 2009. (Image courtesy of Oliver Beauchesne, Science-Metrix.)

for some other form of compensation, usually financial. This arrangement, known as *ghost authorship*, is a problem in the U.S. as well as in other parts of the world. Last year, Shen Yang of Wuhan University in China released estimates that Chinese academics spent more than $145 million on ghostwritten papers in the previous year. That sum is considerable, especially considering reports of low pay to ghostwriters. For example, Associated Press reporter Gillian Wong last year wrote about a Chinese ghostwriter who received the equivalent of $45 per paper for composing professors' research articles. One of our focus-group participants commented that pressures and financial rewards for publications increase Chinese academics' willingness to pay for ghostwriters.

Authors may also remove themselves—either voluntarily or under pressure—from a publication because they fear repercussions for having participated in politically or religiously sensitive research. International collaborators whose research findings may embarrass their governments—for example by exposing weaknesses in health care systems—sometimes ask to be omitted from publications for the sake of their careers.

Undeserved Credit

An omitted author is clearly denied the recognition he or she deserves, but the addition of undeserving authors can also be damaging. Extra names dilute the credit allocated to deserving authors and obscure responsibility for the work. We identify four categories of the added-author problem, distinguishing them according to the motivations for adding an author. These categories overlap to some extent because motivations can be multiple and may not be fully known.

Surprise authorship is when a researcher finds out after publication that his or her name appears on a paper. In some cases, collaborators from different countries do not observe the same practices with regard to coauthorship and review of manuscripts. One scientist told us about a paper published by international colleagues: "I found it by stumbling across the paper in the literature. There it is. This is my name, and there's the paper, and I have never seen this paper."

Gift authorship occurs when someone is given more credit on a paper than he or she deserves. Sometimes a principal investigator decides that it is someone's turn to be on a publication and arranges for that person's name to appear—even if he or she has not done enough to deserve authorship. In other cases, a senior researcher may decide that his collaborators need publications more than he does, so he allocates publication credit generously (or overly generously) to his collaborators. For instance, in some countries it is common practice to include individuals who have only had administrative oversight. One scientist we interviewed complained about too-liberal inclusion standards among his international colleagues: "Sometimes I am very strict about them, basically saying that I'm not going to allow it. Other times I know there is a political reason why they do it, and so unfortunately I may just let it go by." Gift authorship is less benign when it involves an expected *quid pro quo* in the form of future assistance, favors or advantages.

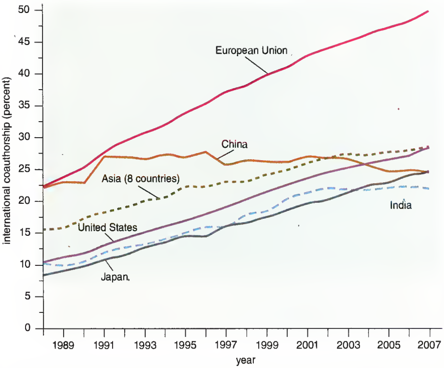
Honorary authorship is often equated with gift authorship, but the motivations are different. Honorary authorship goes to individuals with higher status, as a way of honoring them personally or in their roles as superiors. One scientist told us about working with collaborators in Europe for whom it was standard practice to include "out of respect" the student, the supervisor, the supervisor's mentor and the department chair.

Often, however, the honor bestowed through unearned authorship is not freely given; rather, it is demanded by supervisors, administrators or funders. This issue came up frequently in our interviews and focus groups. An epidemiologist told us,

> People may find that in an international context, if there is a head of the laboratory, that person may expect to go on anything regardless of their contribution or lack of contribution.... So there are issues of sensitivity of where somebody is in terms of the hierarchy.

A focus-group participant explained:

> The culture is different. You may be dealing with a researcher



Articles with authors in two or more countries represent the fastest-growing segment of the science and engineering literature. European Union incentives that encourage collaboration within Europe have had a striking effect. The eight Asian countries represented by the green dotted line are India, Indonesia, Malaysia, the Philippines, Singapore, South Korea, Taiwan and Thailand. (Adapted from data published by the U.S. National Science Board.)

> who in fact has several layers of bosses. And when it comes to negotiating the dollars, they all get involved. When it comes to publications, they all get to have their name on the paper.

In a 2010 *Nature* article, author Karen Kaplan relayed suggestions from academics on how to get tenure—including "name a senior department member as a coauthor on your papers if you're in Europe."

In *legitimizing authorship*, a guest author may be listed because of the credibility that his or her position or status will bring to the publication. As one of our respondents put it,

> People really want to have your name on their paper, sometimes on papers where you ... didn't even know that they were doing the study. But they're using a little bit of the reagent you gave them which, you know, you would give anyone freely.... I think that they want your name on a paper because it may legitimize things.

Another biomedical scientist told us about her experience of being named an author of a paper by a research group in another country, to which she had sent plasmids. She thought the study was done incorrectly, and the findings directly contradicted a paper she had published. She continued,

> They were upset because what they wanted was my name on the paper so that they could submit it to a journal that was a little bit higher up in the hierarchy, because I already had a reputation in the field. So ... this was very uncomfortable because I was saying to them, 'No, your work isn't good enough.' And I was trying to find ways not to say that, but that's frankly what I felt.

In other cases, legitimizing authorship comes into play when an author is added to mask the illegitimate contributions of others. For example, pharmaceutical firms may recruit researchers to serve as figurehead authors of company-authored papers in order to hide potential conflicts of interest. The figurehead is usually paid well to allow his or her name to appear, frequently replacing the actual authors who may be acknowledged in a footnote or may be absent altogether.

Out of Order

Trouble can also arise as collaborators work out the sequence in which their names will appear on a publication. Disciplinary customs differ, particularly in the significance of the first- and last-named authors. Cross-national teams are often cross-disciplinary as well, and the order of authors can be a point of dispute. One of our focus-group members described collaborating with a high-ranking scientist in another country who insisted on taking the last position to signify, he said, his minimal role on the project. In the journal in question, however, the last position indicated a significant responsibility as the corresponding author for the study. The scientist described it as "a very difficult situation." Another scientist we interviewed had worked with a team that included a member of a royal family (a princess) who expected to be the first author on every publication in her country, regardless of her contribution.

Some collaborative teams adjust the order of authors to share credit fairly among themselves. They may adopt alphabetical ordering and rotate the alphabetical list in subsequent papers to take turns being first author. Some teams include a note that all authors, or some subset of the authors, contributed equally. Others take advantage of differences in how contributions are measured in different countries. For example, a biological chemist told us,

> It worked out nicely.... I was offering them either first or last [author position], thinking that they would of course choose last, but they wanted first, which was perfectly fine with me. That put me last, so in that sense, yeah, it was great.

Plagiarism

Plagiarism stakes an implicit authorship claim on someone else's words or ideas. Recent media attention to plagiarism has alerted scientists to the need to check manuscripts for plagiarized material using increasingly sophisticated software.

Plagiarism among international trainees has long been a concern in many countries. In a 2011 article, research ethicists Elizabeth Heitman of Vanderbilt University and Sergio Litewka of the University of Miami Miller School of Medicine discussed numerous factors that may lead to plagiarism. These include the "normalcy of plagiarism" in

some social and cultural environments, vague integrity standards, and rejection of U.S. concepts of originality and intellectual property. Plagiarism may also be a strong temptation for international trainees who have difficulty writing in English but are under pressure to publish in English-language journals. Such problems are not unique to students. In a letter published in *Nature* in 2007, Turkish physicist Ihsan Yilnaz writes,

> For those of us whose mother tongue is not English, using beautiful sentences from other studies on the same subject in our introductions is not unusual.... Borrowing sentences in the part of a paper that simply helps to better introduce the problem should not be seen as plagiarism. Even if our introductions are not entirely original, our results are.

Cultural perspectives on authority also influence writers' attitudes toward plagiarism. One of our interviewees said, "It took me some years to figure out that there's an idea at large that anything that comes from authority figures is held in high regard." She noted that material in the published literature is assumed to come from a person of authority, so students may conclude, "What right do I have to change that person's words?"

Ready access to articles online has made plagiarism easier. Environmental engineer James Leckie observed in an interview that students in some cultures do not believe they are doing anything wrong when they plagiarize. They argue that "if the authors didn't want you to use their material, they wouldn't publish it, but since it's published and accessible, it should be free for everyone to use in any way."

A specific issue in the global context is *translation plagiarism*—that is, translating a publication in whole or in part and publishing the translation as one's own work without acknowledging the original authors. One U.S. biomedical researcher we interviewed encountered this problem with a collaborator in a developing country. The collaborator asked him to review and endorse a book manuscript, which turned out to be a translation of materials that the U.S. researcher had given to the collaborator, for which the collaborator intended to take sole credit. The U.S. scientist saw that the country in question would derive great value from having the materials in its native language, but he confronted the collaborator about the plagiarism:

> I said, 'I've got a real problem with this, because I recognize why you've done it and all of this, but the fact is that you have put your name on other people's words, and what you did is you translated it.' And he said, 'Well yes!... What's the problem with that? And aren't I your friend? And look at what I have done for you!' and things like that. It was pretty awkward.

Skills for the Global Context

Authorship sits at the intersection between collective effort and individual ambition. Scientists participate in international collaborations for many reasons, including a belief that collaboration will benefit all involved. But the pursuit of individual recognition cannot be completely eliminated. Collaborators must pay careful attention to authorship in order to share credit and responsibility fairly among all team members.

The most helpful way to deal with authorship issues is to agree on general principles for authorship at the beginning of the collaboration and then to agree on authorship of each article when its content is first outlined. A U.S. scientist who is experienced at international collaboration told us, "You have to have the guts to tackle [authorship issues] before you go into it." Another said,

> I found it difficult at first. But it was very clear that if there is any doubt as to how authorship—especially credit—is going to be divvied up, it is better to approach that before, rather than after, just so everybody has a pretty good idea.

It is clear, though, that discussions alone will not clear away all authorship problems. Several resources provide additional guidance. Collaborators should consult the policies of the journals to which they plan to submit their work. International guidance is available through the the Council of Science Editors, the Committee on Publication Ethics and the International Committee on Medical Journal Editors' Uniform Requirements for Manuscripts Submitted to Biomedical Journals. The *Singapore Statement*, released in conjunction with the Second World Conference on Research Integrity in 2010, provides a succinct statement that responsible authorship is a duty of researchers worldwide. Scientists can join AuthorAID, a free, international research community that supports researchers from developing countries with services such as networking and mentorship.

The U.S. scientists we interviewed were aware of the ethical complexity of authorship and the importance of meeting high standards in practice. But they were also careful to take their collaborators' perspectives into account. A focus-group participant said

> the problem is that our ignorance of the way these kinds of systems work in other countries can sometimes be really detrimental to the way the research is performed and expressed and published.

Leckie provides a cautionary example, drawn from his own experience working with an Asian collaborator:

> I once was designing a research activity with this fellow, and he brought to me very early on several drafts of his part of the proposal. And on it, he had the name of a division head who was associated with the overall program but had nothing to do with his research activity. The division head had no expertise relevant to the proposal and was not going to contribute anything. I told my collaborator, 'Look, that guy's not doing anything. Take his name off.' And he said, 'Well, I can't.' Then I said, 'Well then, take my name off.' And so we had a real confrontation, and in the end my collaborator took the name off, and it resulted in an attempt to fire him.

Maintaining the integrity of authorship is complicated in the global context, but the stakes can be high for all concerned. It is worth the time and effort required to get it right.

Bibliography

Anderson, Melissa S., and Nicholas H. Steneck, eds. 2010. *International Research Collaborations: Much to be Gained, Many Ways to Get in Trouble*. New York: Routledge.

Heitman, Elizabeth, and Sergio Litewka. 2011. International perspectives on plagiarism and considerations for teaching international trainees. *Urologic Oncology* 29:104–108.

Martinson, Brian C., Melissa S. Anderson, and Raymond De Vries. 2005. Scientists behaving badly. *Nature* 435: 737–738.

Wong, Gillian. 2010. In China, academic cheating is rampant. *The Boston Globe* (April 11).

Science Observer

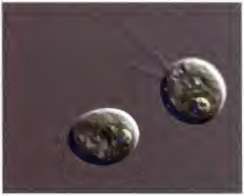
Dying Generously

For some single-celled organisms, self-destruction is neighborly

Suicide is an evolutionary conundrum. Single-celled organisms regularly kill themselves in reaction to stresses they might have survived, but it's not obvious why natural selection permits such volatile behavior.

One explanation, that suicide can benefit neighbors and relatives of the deceased cell, got a boost with new experiments published in the February issue of *American Naturalist*. Working together at the University of Arizona, Pierre Durand, Armin Rashidi and Richard Michod found that when the unicellular alga *Chlamydomonas reinhardtii* dies by suicide, it releases compounds that help surviving cells grow faster.

In multicellular organisms, programmed cell death (PCD) demands no special explanation because it's not really suicide. The whole organism has to survive and reproduce, and its component cells should do whatever it takes to help. In humans, for example, PCD shapes fingers, blood vessels and other anatomical features during development, and eliminates damaged cells throughout life.



If these *Chlamydomonas reinhardtii* cells were to die suddenly, they would spill toxic contents into their environment and harm nearby cells. But when they die by suicide, or programmed cell death, they instead release compounds that help surviving cells multiply faster. (Photograph by Deborah Bock, courtesy of Pierre Durand.)

When an individual cell is also an entire organism, however, PCD is counterintuitive—so much so that when biologists discovered suicide in marine algae, they were met with skepticism. "For the longest time, people didn't even believe that this pathway existed in these organisms because they were unicellular," says Kay Bidle, a molecular ecologist who studies marine algae at Rutgers University.

But the discovery was no fluke. Several years later, researchers have found PCD everywhere they've looked. Fungi, protozoans, red and green algae, and even bacteria have genetic programs that, when activated, orchestrate self-destruction. Suicidal cells actively expend energy to shrink, chop up their own DNA and engineer other fatal changes.

Although premature death has turned out to be common in single-celled organisms, biologists still struggle to explain it. "If there's a chance of you surviving you should do it," says Michod. But if in death, a cell could benefit other individuals that share its genes, natural selection might keep suicidal traits in the population.

To test this hypothesis, the team grew three clonal cultures of *C. reinhardtii*. They plunged one flask of algae into a water bath of 50 degrees Celsius for 10 minutes. That's enough to start scalding human skin and to launch *C. reinhardtii's* death program. Within less than a day, those cells had all perished, and showed telltale signs of suicide, such as chopped-up DNA.

The second culture died by necrosis—that is, sudden "unplanned" death during which cells dump their contents into the environment unaltered. To simulate necrosis, the team used a sonicator probe to tear the algae apart with high-pitched sound waves. Michod acknowledges that sonication was an artificial way of killing the cells. But, he says, it was the only assault they could conceive that would kill the cells before they could even start PCD.

The third culture, the control, enjoyed a normal unstressed life for another 18 hours. Then the team supplemented new *C. reinhardtii* cultures with the liquid in which suicide, necrosis, or business as usual had transpired. In a final control, some cultures got fresh nutrient broth with no supplement. "It was very dramatic," Michod says. Over dozens of replicate experiments, the leftovers from PCD made cells grow faster than did fresh nutrient broth or leftovers from the control culture. Depending on the concentration, the PCD supplement could double a culture's growth over four days.

The broth contaminated by necrosis, on the other hand, was a potent growth inhibitor. "Cells carry a lot of toxic things," Michod says. If those contents, such as chlorophyll, enter the environment unaltered, they poison the neighbors. Still, it's not clear how PCD rendered chlorophyll harmless and liberated growth-enhancing compounds. That, says lead author Durand, is the subject of his current studies at the University of Witwatersrand in South Africa. He also wants to know how specific the benefit is: The kin-selection argument would be strengthened if the products of PCD are helpful only to members of the same species or strain.

The results aren't the first to suggest that suicide can aid neighboring cells. Related phenomena have turned up in ultraviolet-stressed *C. reinhardtii* and in aging yeast. But they are among the first to show a direct benefit to surviving cells and to compare the effects of different ways of dying. Frank Madeo, who studies PCD in yeast at the University of Graz in Aus-

tria, says the new study strengthens the idea that natural selection favors PCD because it helps the whole group of cells. "The overall story is nice," he says—although he wants to know more details about exactly how the beneficial effect works. He points out that an apparent boost in growth could actually be a decreased death rate in the culture.

At the same time, the new results leave intact several other hypotheses about the benefits of PCD. For instance, suicide could be an accidental by-product of some metabolic process—such as adapting to milder stresses—that makes cells competitive earlier in life. Or it may help algae combat viral infections. If virus-infected cells self-destruct before the pathogen can replicate, they spare their neighbors the infection. Bidle predicts all these factors are probably "part of the story" in every lineage of unicellular organisms.—*Elsa Youngsteadt*

Making Better Maps of Food Deserts

Neighborhoods with little or no access to healthful food can be located and studied using GIS mapping

Food deserts—areas where consistent access to healthful food is limited or nonexistent—have gained attention recently. Michelle Obama has launched an initiative to end them in the United States by 2017. But in many areas, we still lack comprehensive, fine-grain knowledge of where food access is most challenging.

"It's something that kind of snuck up on us as a scientific community," says Kirk Goldsberry, a professor in the Department of Geography at Michigan State University. Goldsberry and his colleagues Chris Duvall and Phil Howard, also of MSU, have made a new series of maps showing access to fresh produce in the Lansing, Michigan, area.

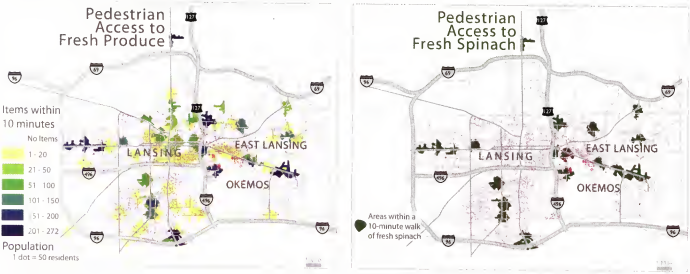
The team wanted to measure food access using a deeper level of data than is often available. So, Goldsberry says, they visited every fresh-produce retailer in the area, locating 447 distinct produce items for sale. Then, using GIS (Geographic Information Systems), they mapped where each item could be bought.

Next they estimated routes and times that shoppers, both pedestrians and drivers, might take to the store, using a 10-minute walk (with an average speed of 3 miles per hour) and a 10-minute drive as their respective limits. Putting all the layers together yielded an interactive atlas of fresh-produce accessibility.

Jason Gilliland, a geographer at the University of Western Ontario who studies food deserts, likes the map's design. "It's one of the most visually appealing food-desert maps in terms of cartography," he says. The atlas is already being used, says Goldsberry, to determine where to place community gardens. He hopes it will inform other community action and policy-making decisions, and that its techniques will be replicated. "We hope that people in Denver, Colorado, or people in Austin, Texas, see this map and say, 'I want to see how my city looks through this lens,'" he says.

"Standardization of methods is critical for comparing results," says Gilliland. Using higher-resolution data, as Goldsberry's team has, is a good start. But Gilliland notes that we need better studies to help define map parameters. How long does the average pedestrian commute to a food source actually take, for instance? Alternately, maps could show availability in 100-meter increments rather than using a time limit. "Then you'd have a true accessibility surface," he says.

And there are other factors competing for space on such maps. Something Goldsberry would like to add is change in availability of produce over time. In addition, maps would ideally reflect all the transit modes and food sources people use. So public-transit and bicycling data are valuable. Existing community gardens and food banks are important access points to include, Gilliland notes. And getting help from community members with data gathering can improve accuracy.



A map made by geographers at Michigan State University shows areas of good and poor access to fresh produce in the Lansing, Michigan, area. The map at left is made up of data from 447 individual maps of access to produce items, like the map for spinach shown at right. Pink dots in the main map represent areas of denser population. The interactive atlas is available at http://geo.msu.edu/food/media/atlas/atlas.htm.

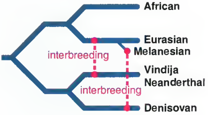
Then there are things that may not even be mappable, but that nonetheless affect people's lives. "Much more goes into food access than can ever be captured by a GIS map," says Tim Stallmann, a Ph.D. student in geography at the University of North Carolina and a member of the Counter Cartographies Collective. "There's a whole dimension around money and differential access to stores, how different stores make different groups of people feel welcomed. There's the amount of time folks have to go shopping in the first place, and when they have it."

As geographers work to refine maps of food deserts, what can be done about the problem itself? A three-year study by Gilliland and Kristian Larsen, published in 2009 in the journal *Health and Place*, documented a 12-percent drop in average prices for healthful foods in an underserved neighborhood in London, Ontario, after a farmer's market opened there. This decrease occurred even as prices at supermarkets in London rose by 9 percent.

Increasing access to healthful food could have dramatic results for public health. As Gilliland puts it, "We know you are what you eat, and that includes, you are *where* you eat."—*Anna Lena Phillips*

Paleogenomic Puzzles

DNA sequences of extinct hominins could rewrite human ancestry



If recent genomic studies are correct, some modern people can claim Neanderthals and a newly identified hominin group called Denisovans as distant ancestors. (Chart adapted from Pääbo et al. 2010. *Nature*)

A research team led by Svante Pääbo of the Max Planck Institute for Evolutionary Anthropology dropped two bombshells on the field of paleoanthropology last year. But how the new data may rewrite our understanding of human ancestry remains uncertain today.

After sequencing multiple genomes of Neanderthals, our extinct but closest known relatives, the researchers concluded that *Homo sapiens* early to reach Eurasia from Africa had interbred with Neanderthals. The evidence: Non-Africans living today share distinctive DNA sequence with Neanderthals, who lived in Eurasia from 200,000 to 28,000 years ago, they say

Later, they identified a new group of extinct hominins—called Denisovans—using DNA retrieved from a young female's finger bone recovered in Denisova Cave in Siberia. Denisovans, they say, also interbred with *Homo sapiens*, but only a small group of modern people—those living today in Papua New Guinea—are known to retain evidence of that ancestry in their genomes.

"The situation is fantastic and a bit of a first in genetics," says Nick Patterson, a computational biologist at the Broad Institute in Cambridge, Massachusetts who is closely involved in the research. "The normal thing is that you have a species, maybe even an ancient species. But the physical characteristics are very well known. Then you learn about the genetics. "

Not all paleoanthropologists are convinced by these findings. Richard Potts, director of the human origins program at the Smithsonian Institution, says the Neanderthal-human interbreeding evidence was highly convincing, in isolation. But when the Denisovan argument emerged later, he wondered if the scientists had found something other than what they concluded. He suspects the sampled Neanderthal, Denisovan and *Homo sapien* genomes could retain DNA sequences from an earlier hominin, possibly *Homo heidelbergensis*.

"What we may be seeing instead is the parceling and division of DNA from a last common ancestor. That could have a different echo in different groups of modern humans," Potts says.

If *Homo sapiens* dispersing from Africa did interbreed with hominins called Denisovans, their offspring appear to have been among people who kept migrating east and south, down to Melanesia.

Pääbo's team is trying to strengthen evidence for their findings. Patterson, for one, is working on what he calls a genetic recombination clock to better estimate when *Homo sapiens* and Neanderthals might have interbred. Pääbo says they are also analyzing more bones from the Denisova Cave and from other locations in the hope that they will yield more DNA. "It would be interesting to see how widespread Denisovans were in the past," he says.

Pääbo, director of evolutionary genetics at the Max Planck Institute in Leipzig, hopes that DNA can also be retrieved from *Homo sapiens* who reached Eurasia during our species' first sizable dispersal from Africa about 80,000 years ago. Potts of the Smithsonian sees promise in the same line of research. He is hopeful that DNA someday can be recovered specifically from the Qafzeh fossils from Israel, which are about 130,000 years old, and from the Liujiang fossil from China, which is at least 68,000 years old. Sequences from *H. heidelbergensis*, he says, could be awfully interesting too.

"Getting genes out of fossil bone is terrific. It opens new possibilities," he says.—*Catherine Clabby*

This roundup summarizes some notable recent items about scientific research, selected from news reports compiled in Sigma Xi's free electronic newsletters Science in the News Daily *and* Science in the News Weekly. *Online:* ***http://sitn.sigmaxi.org*** *and* ***http://www.americanscientist.org/sitnweekly***

Grapes Need Sex

Wine and table grapes *(Vitis vinifera vinifera)* haven't had much recombination in the past 7,000 years. To keep specific flavors and traits together, farmers have grown grapes from cuttings for millenia. That clonal reproduction maintains genetic diversity but does not shuffle it. In an analysis of 950 grape samples from the U.S. Department of Agriculture collection, more than half the accessions were clones of others: Seventeen different pinots, for example, were all twins save for a few mutations. Encouraging some grape sex might let growers incorporate traits for pest resistance from rare and wild varieties. Using modern genetic markers, farmers could still track the flavors and colors required for flawless wine.

Myles, S., et al. Genetic structure and domestication history of the grape. Proceedings of the National Academy of Sciences *108:3530–3535 (March 1)*

Methane Monsoon

On Saturn's moon Titan, the poles are dotted with hydrocarbon seas but the tropics are an arid expanse of dunes and channels. According to new images from the Cassini spacecraft, that desert may be seasonally quenched with methane rain. In late September 2010, an expansive cloud system crossed Titan's equator and left behind a long dark stain on 500,000 square kilometers of the moon's surface. Researchers think the discolored ground was wet or flooded, and began to drain or evaporate within a month. If the spring rains continue, they may help explain Titan's desert stream beds.

Turtle, E. P., et al. Rapid and extensive surface changes near Titan's equator: Evidence of April showers. Science, *331:1414–1417 (March 18)*

Mix and Match Genes

Bacteria and archaea have two ways of gaining new metabolic abilities: Extra copies of their own genes can evolve novel functions, and they can acquire genes from other species. Biologists thought the first method—predominant in eukaryotes—was also the most important in prokaryotes. But a new study of 110 genomes in eight microbial lineages suggests the reverse. When those organisms gained new versions of genes, they did it by acquisition 88 percent of the time. That's probably because copies of existing genes may be deleted as excess baggage before they evolve useful mutations. New genes from another species, however, can serve a novel function right away.

Treangen, T. J. and E. P. C. Rocha. Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. PLoS Genetics *7: e1001284 (January 27)*

Shivering in Europe

Early humans lived in Europe for 700,000 long cold years before they mastered fire. That conclusion comes as a surprise to many archaeologists, who thought that controlled flames must have been a prerequisite for human migration into colder climates. But a review of data from more than 100 archaeological sites across Europe revealed no evidence of habitual fire use until 400,000 years ago. After that, remains of ancient hearths and scorched tools appeared with increasing frequency. The authors suggest that the first humans in cold climates must have relied on high metabolism—stoked with lots of protein and an active lifestyle—to weather frosty boreal winters.

Roebroeks, W., and P. Villa. On the earliest evidence for habitual use of fire in Europe. Proceedings of the National Academy of Sciences *(published online March 14)*

Fire-Ant Headquarters

In the 1930s, fire ants *(Solenopsis invicta)* from Argentina set up shop in Alabama and spread quickly through the southern United States. Now, the insects' new home serves as a launch pad for more invasions around the globe. An analysis of fire-ant DNA from 2,144 colonies indicated that California, China, Taiwan and Australia experienced eight distinct introductions from the southern U.S. A ninth ant envoy passed through California before continuing to Taiwan. Researchers aren't sure why the U.S. fire ants are such super-invaders, but the same traits that helped the first populations settle in the South may give them an edge elsewhere too.

Ascunce, M. S., et al. Global invasion history of the fire ant Solenopsis invicta. Science *331:1066–1068 (February 25)*

Amoebas Save Seed

Even single-celled organisms can farm. Social amoebas *(Dictyostelium discoideum)* live alone and prey upon bacteria until food becomes scarce. Then they form a slug-like collective and creep away together to release spores. Working with 35 strains of wild *D. discoideum*, researchers discovered that about one-third of the strains always took uneaten bacteria with them when they crawled off. The bacteria then dispersed along with spores, seeding new habitat with good food. That gave the farming strains an advantage when their spores landed on sterile media or on soil with inedible bacteria. But where food was already abundant, the nonfarmers reproduced faster. A patchy food landscape probably maintains the balance between farming and nonfarming strains in nature.

Brock, D. A., et al. Primitive agriculture in a social amoeba. Nature *469:393–396 (January 20)*

Bird-Inspired Design

Woodpecker anatomy has inspired a new shock absorber for delicate electronics in plane crashes and bombings. The birds incessantly slam their heads against hard surfaces without brain damage. So engineers studied X-ray scans of a golden-fronted woodpecker *(Melanerpes albifrons)* and modeled its shock-absorbing structures—such as its elastic beak and partially spongy skull. They mimicked key parts of the bird's head with a cylinder made of metal, rubber and porous glass. Diodes and other delicate devices inside the cylinder survived more than 99 percent of the time when they hit a wall with a deceleration 60,000 times that of gravity. Survival was only 74 percent in conventional shock absorbers.

Yoon, S.-H., and S. Park. A mechanical analysis of woodpecker drumming and its application to shock-absorbing systems. Bioinspiration and Biomimetics *6:016003 (January 17)*

Global Energy: The Latest Infatuations

In energy matters, what goes around, comes around —but perhaps should go away

Vaclav Smil

To follow global energy affairs is to have a never-ending encounter with new infatuations. Fifty years ago media ignored crude oil (a barrel went for little more than a dollar). Instead the western utilities were preoccupied with the annual double-digit growth of electricity demand that was to last indefinitely, and many of them decided that only large-scale development of nuclear fission, to be eventually transformed into a widespread adoption of fast breeder reactors, could secure electricity's future. Two decades later, in the midst of the second energy "crisis" (1979–1981, precipitated by Khomeini's takeover of Iran), rising crude oil prices became the world's prime existential concern, growth of electricity demand had slumped to low single digits, France was the only nation that was seriously pursuing a nuclear future, and small cars were in vogue.

After world crude oil prices collapsed in 1985 (temporarily below $5 per barrel), American SUVs began their rapid diffusion that culminated in using the Hummer H1, a civilian version of a U.S. military assault vehicle weighing nearly 3.5 tonnes, for trips to grocery stores—and the multinational oil companies were the worst performing class of stocks of the 1990s.

Vaclav Smil is a Distinguished Professor at the University of Manitoba. His interdisciplinary research deals with interactions of energy, environment, food, economy, population and technical advances. He is the author of 30 books on these topics; the latest ones (in 2010) include Energy Myths and Realities, Energy Transitions *and* Two Prime Movers of Globalization. *E-mail: vsmil@cc.umanitoba.ca Website: http://www.vaclavsmil.com*

The first decade of the 21st century changed all that, with constant fears of an imminent peak of global oil extraction (in some versions amounting to nothing less than lights out for western civilization), catastrophic consequences of fossil fuel-induced global warming and a grand unraveling of the post-WW II world order.

All of this has prompted incessant calls for the world to innovate its way into a brighter energy future, a quest that has engendered serial infatuations with new, supposedly perfect solutions: Driving was to be transformed first by biofuels, then by fuel cells and hydrogen, then by hybrid cars, and now it is the electrics (Volt, Tesla, Nissan) and their promoters (Shai Agassi, Elon Musk, Carlos Ghosn) that command media attention; electricity generation was to be decarbonized either by a nuclear renaissance or by ubiquitous wind turbines (even Boone Pickens, a veteran Texas oilman, succumbed to that call of the wind), while others foresaw a comfortable future for fossil fuels once their visions of mass carbon capture and sequestration (CCS) were put in practice. And if everything fails, then geoengineering—manipulating the Earth's climate with shades in space, mist-spewing ships or high-altitude flights disgorging sulfur compounds—will save us by cooling the warming planet.

This all brings to mind Lemuel Gulliver's visit to the grand academy of Lagado: No fewer than 500 projects were going on there at once, always with anticipation of an imminent success, much as the inventor who "has been eight years upon a project for extracting sunbeams out of cucumbers" believed that "in eight years more, he should be able to supply the governor's gardens with sunshine, at a reasonable rate"—but also always with complaints about stock being low and entreaties to "give … something as an encouragement to ingenuity." Admittedly, ideas for new energy salvations do not currently top 500, but their spatial extent puts Lagado's inventors to shame: Passionately advocated solutions range from extracting work from that meager 20-Kelvin difference between the surface and deep waters in tropical seas (OTEC: ocean thermal energy conversion) to Moon-based solar photovoltaics with electricity beamed to the Earth by microwaves and received by giant antennas.

And continuous hopes for success (at a low price) in eight more years are as fervent now as they were in the fictional 18th century Lagado. There has been an endless procession of such claims on behalf of inexpensive, market-conquering solutions, be they fuel cells or cellulosic ethanol, fast breeder reactors or tethered wind turbines. And

Figure 1. In *Gulliver's Travels* Lemuel visits the grand academy of Lagado, where more than 500 research projects were ongoing, always with expectations of an early payoff. The inventor here, as shown in an illustration by Milo Winter in the 1930 edition, is soon to extract sunbeams from cucumbers. The author sees an analogy with the current attitude toward energy in the United States and Canada. Instead of facing the fact that these two countries use twice as much energy per capita as other wealthy nations with similar indicators of human development and instead of learning to do as much with much less, these countries continue to search for technical fixes to maintain, even to increase, their use of energy.


EXT
CUC
MILO WINTER

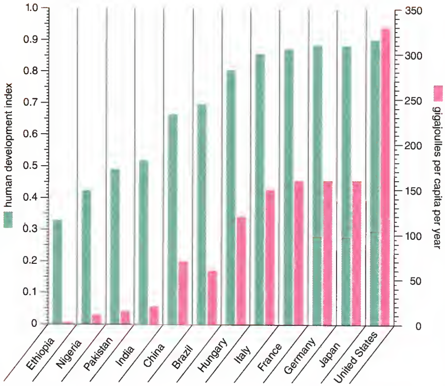


Figure 2. At very low and low per capita consumption levels, higher use of energy is clearly tied to rising index of human development, but once energy per capita reaches about 150 gigajoules per year, the correlation breaks down. More is not better.

energy research can never get enough money to satisfy its promoters: In 2010 the U.S. President's council of advisors recommended raising the total for U.S. energy research to $16 billion a year; that is actually too little considering the magnitude of the challenge—but too much when taking into account the astonishing unwillingness to adopt many readily available and highly effective existing fixes in the first place.

Enough to Go Around?

Although all this might be dismissed as an inevitable result of the desirably far-flung (and hence inherently inefficient) search for solutions, as an expected bias of promoters devoted to their singular ideas and unavoidably jockeying for limited funds, I see more fundamental, and hence much more worrisome, problems. Global energy perspective makes two things clear: Most of humanity needs to consume a great deal more energy in order to experience reasonably healthy lives and to enjoy at least a modicum of prosperity; in contrast, affluent nations in general, and the United States and Canada in particular, should reduce their excessive energy use. While the first conclusion seems obvious, many find the second one wrong or outright objectionable.

In 2009 I wrote that, in order to retain its global role and its economic stature, the United States should

> provide a globally appealing example of a policy that would simultaneously promote its capacity to innovate, strengthen its economy by putting it on sounder fiscal foundations, and help to improve Earth's environment. Its excessively high per-capita energy use has done the very opposite, and it has been a bad bargain because its consumption overindulgence has created an enormous economic drain on the country's increasingly limited financial resources without making the nation more safe and without delivering a quality of life superior to that of other affluent nations.

I knew that this would be considered a nonstarter in the U.S. energy policy debate: Any calls for restraint or reduction of North American energy use are still met with rejection (if not derision)—but I see that quest to be more desirable than ever. The United States and Canada are the only two major economies whose average annual per capita energy use surpasses 300 gigajoules (an equivalent of nearly 8 tonnes, or more than 50 barrels, of crude oil). This is twice the average in the richest European Union (E.U.) economies (as well as in Japan)—but, obviously, Pittsburghers or Angelenos are not twice as rich, twice as healthy, twice as educated, twice as secure or twice as happy as inhabitants of Bordeaux or Berlin. And even a multiple adjustment of national per capita rates for differences in climate, typical travel distances and economic structure leaves most of the U.S.–E.U. gap intact: This is not surprising once it is realized that Berlin has more degree heating days than Washington D.C., that red peppers travel the same distance in refrigerated trucks from Andalusia to Helsinki as they do from California's Central Valley to Illinois, and that German exports of energy-intensive machinery and transport-equipment products surpass, even in absolute terms, U.S. sales.

Moreover, those who insist on the necessity and desirability of further growth of America's per capita energy use perhaps do not realize that, for a variety of reasons, a plateau has been reached already and that (again for many reasons) any upward departures

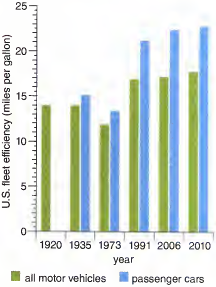


Figure 3. Although the efficiency of internal combustion engines has increased substantially in the past 90 years (particularly when the adoption of diesel-powered cars is taken into account), the average performance of motor vehicles in the Unites States has improved only from about 14 miles per gallon to about 18 mpg.

are highly unlikely. In 2010 U.S. energy consumption averaged about 330 gigajoules per capita, nearly 4 percent lower than in 1970, and even the 2007 (pre-crisis) rate of 355 gigajoules (GJ) per capita was below the 1980 mean of 359 GJ. This means that the U.S. per capita consumption of primary energy has remained essentially flat for more than one generation (as has British energy use). How much lower it could have been can be illustrated by focusing on a key consumption sector, passenger transport.

Planes, Trains and Automobiles

After 1985 the United States froze any further improvements in its corporate automobile fuel efficiency (CAFE), encouraged a massive diffusion of exceptionally inefficient SUVs and, at the same time, failed to follow the rest of modernizing world in building fast train links. For 40 years the average performance of the U.S. car fleet ran against the universal trend of improving efficiencies: By 1974 it was lower (at 13.4 miles per gallon [mpg]) than during the mid-1930s! Then the CAFE standards had doubled the efficiency of new passenger cars by 1985, but with those standards subsequently frozen and with the influx of SUVs, vans and light trucks, the average performance of the entire (two-axle, four-wheel) car fleet was less than 26 mpg in 2006 or no better than in 1986—while a combination of continued CAFE upgrades, diffusion of new ultra low-emission diesels (inherently at least 25–30 percent more efficient than gasoline-powered cars) and an early introduction of hybrid drives could have raised it easily to more than 35 or even 40 mpg, massively cutting the U.S. crude oil imports for which the country paid $1.5 trillion during the first decade of the 21st century.

And the argument that its large territory and low population density prevents the United States from joining a growing list of countries with rapid trains (traveling 250–300 kilometers per hour or more) is wrong. The northeastern megalopolis (Boston-Washington) contains more than 50 million people with average population density of about 360 per square kilometer and with nearly a dozen major cities arrayed along a relatively narrow and less than 700-kilometer long coastal corridor. Why is that region less suited to a rapid rail link than France, the pioneer of European rapid rail transport, with a population of 65 million and nationwide density of only about 120 people per square kilometer whose *trains à grande vitesse* must radiate from its capital in order to reach the farthest domestic destinations more than 900 kilometers away? Apparently, Americans prefer painful trips to airports, TSA searches and delayed shuttle flights to going from downtown to downtown at 300 kilometers per hour.

In a rational world animated by rewarding long-term policies, not only the United States and Canada but also the European Union should be boasting about gradual reductions in per capita energy use. In contrast, modernizing countries of Asia, Latin America and, most of all, Africa lag so far behind that even if they were to rely on the most advanced conversions they

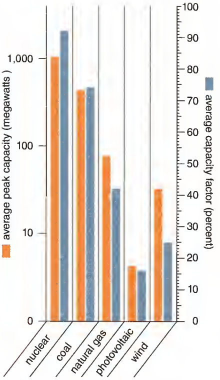


Figure 4. Generation of electricity by wind turbines and photovoltaic (PV) cells differs in two fundamental ways from thermal electricity production. First, as shown in the left column, average capacities of photovoltaic and wind farms are smaller than those of nuclear, coal and even natural gas-powered generators. (Note the logarithmic scale.) Second, the percentage of time that the generators can work at full capacity (load factor) is much lower. (The capacity factor of gas-fired generators is constrained not by their ability to stay online but by their frequent use as intermittent sources to meet demand peaks.) Moreover, differences in capacity factors will always remain large. In 2009 the load factor averaged 74 percent for U.S. coal-fired stations, and the nuclear ones reached 92 percent, whereas wind turbines managed only about 25 percent. (All plots show the U.S. averages in 2009.)

Figure 5. Canada's Sarnia Photovoltaic Power Plant became the world's largest PV plant at 80 megawatts of peak power when it was completed in September 2010. It consists of about 1.3 million thin-film PV panels covering about 966,000 square meters, but its capacity factor is expected to be only about 17 percent. (Photo courtesy of First Solar.)

would still need to at least quadruple (in India's case, starting from about 20 GJ per capita in 2010) their per capita supply of primary energy or increase their use by more than an order of magnitude—Ethiopia now consumes modern energies at a rate of less than 2 GJ per capita—before getting to the threshold of a decent living standard for most of their people and before reducing their huge internal economic disparities.

China has traveled further, and faster, along this road than any other modernizing nation. In 1976 (the year of Mao Zedong's death) its average per capita energy consumption was less than 20 GJ per capita, in 1990 (after the first decade of Deng Xiaoping's modernization) it was still below 25 GJ, and a decade later it had just surpassed 30 GJ per capita. By 2005 the rate had approached 55 GJ and in 2010 it reached 70 or as much as some poorer E.U. countries were consuming during the 1970s. Although China has become a major importer of crude oil (now the world's second largest, surpassed only by the United States) and it will soon be importing large volumes of liquefied natural gas and has pursued a large-scale program of developing its huge hydrogenation potential, most of its consumption gains have come from an unprecedented expansion of coal extraction. While the U.S. annual coal output is yet to reach one billion tonnes, China's raw coal extraction rose by one billion tonnes in just four years between 2001 and 2005 and by nearly another billion tonnes by 2010 to reach the annual output of 3 billion tonnes.

China's (and, to a lesser degree, India's) coal surge and a strong overall energy demand in Asia and the Middle East have been the main reason for recent rises of CO_2 emissions: China became the world's largest emitter in 2006, and (after a small, economic crisis-induced, decline of 1.3 percent in 2009) the global total of fossil fuel-derived CO_2 emissions set another record in 2010, surpassing 32 billion tonnes a year (with China responsible for about 24 percent). When potential energy consumption increases needed by low-income countries are considered together with an obvious lack of any meaningful progress in reducing the emissions through internationally binding agreements (see the sequential failures of Kyōtō, Bali, Copenhagen and Cancún gatherings), it is hardly surprising that technical fixes appear to be, more than ever, the best solution to minimize future rise of tropospheric temperatures.

Renewable Renaissance?

Unfortunately, this has led to exaggerated expectations rather than to realistic appraisals. This is true even after excluding what might be termed zealous sectarian infatuations with those renewable conversions whose

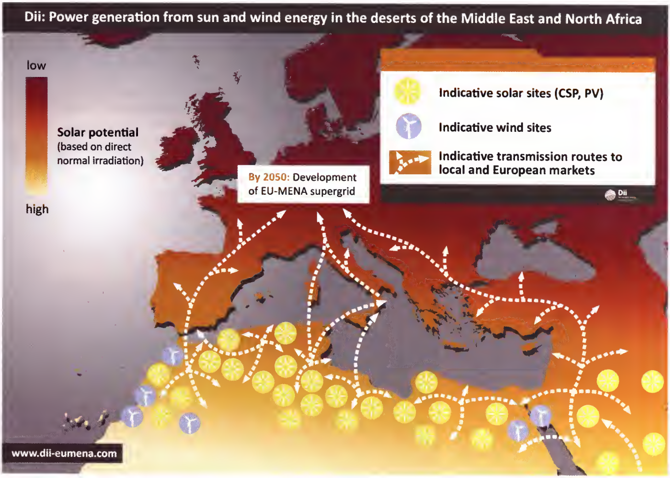


Figure 6. Desertec is perhaps the most ambitious renewable energy plan yet conceived. Most of its electricity would come from concentrating solar thermal power plants in desert regions of northern Africa and the Middle East and would be transmitted by intercontinental high-voltage direct-current lines. The scale of the challenge is obvious, and recent political upheavals across the entire region where these conversions were to take place are not encouraging. (Map courtesy Desertec: http://www.dii-eumena.com/home.html)

limited, exceedingly diffuse or hard-to-capture resources (be they jet stream winds or ocean waves) prevent them from becoming meaningful economic players during the next few decades. Promoters of new renewable energy conversions that now appear to have the best prospects to make significant near-term contributions—modern biofuels (ethanol and biodiesel) and wind and solar electricity generation—do not give sufficient weight to important physical realities concerning the global shift away from fossil fuels: to the scale of the required transformation, to its likely duration, to the unit capacities of new convertors, and to enormous infrastructural requirements resulting from the inherently low power densities with which we can harvest renewable energy flows and to their immutable stochasticity.

The scale of the required transition is immense. Ours remains an overwhelmingly fossil-fueled civilization: In 2009 it derived 88 percent of its modern energies (leaving traditional biomass fuels, wood and crop residues aside) from oil, coal and natural gas whose global market shares are now surprisingly close at, respectively, 35, 29 and 24 percent. Annual combustion of these fuels has now reached 10 billion tonnes of oil equivalent or about 420 exajoules (420×10^{18} joules). This is an annual fossil fuel flux nearly 20 times larger than at the beginning of the 20th century, when the epochal transition from biomass fuels had just passed its pivotal point (coal and oil began to account for more than half of the global energy supply sometime during the late 1890s).

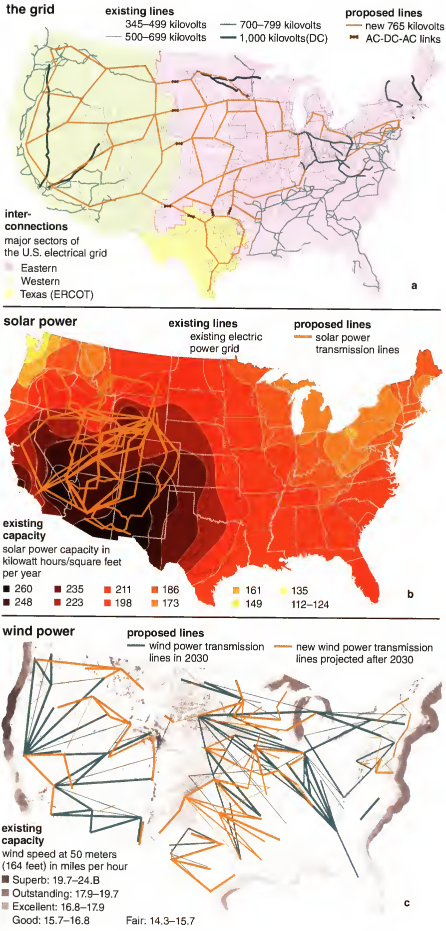


Figure 7. In the United States the foremost problem in replacing much conventional electrical production with renewables is to get power from where it is most efficiently produced to where it is most needed. The existing U.S. grid is divided into zones *(a)*, which do not normally share power on a large scale, and a new nationwide grid would be needed to connect them. The availability of solar power is concentrated in the southwestern United States *(b)*. Even there, many new transmission lines would be required to transfer PV electricity from production to demand. Challenges of delivering wind-generated electricity from the Great Plains and other windy regions to the coastal concentrations of heaviest demand are even greater *(c)*. (Maps adapted from National Public Radio: http://www.npr.org/templates/story/story.php?storyId=110997398)

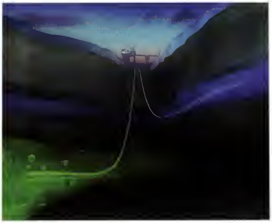


Figure 8. Since 1996 Statoil, the Norwegian state oil company, has been capturing annually one million tonnes of carbon dioxide from natural gas production at Sleipner West field (250 kilometers offshore in the North Sea) and sequestering it in an aquifer more than 800 meters below the seabed. But the global challenge is three to four orders of magnitude greater: If sequestration were to slow down the rise of atmospheric CO_2 it would have to proceed at an annual rate of many billions of tonnes. (Image cuourtesy of Statoil.)

Energy transitions—shifts from a dominant source (or a combination of sources) of energy to a new supply arrangement, or from a dominant prime mover to a new converter—are inherently prolonged affairs whose duration is measured in decades or generations, not in years. The latest shift of worldwide energy supply, from coal and oil to natural gas, illustrates how the gradual pace of transitions is dictated by the necessity to secure sufficient resources, to develop requisite infrastructures and to achieve competitive costs: It took natural gas about 60 years since the beginning of its commercial extraction (in the early 1870s) to reach 5 percent of the global energy market, and then another 55 years to account for 25 percent of all primary energy supply. Time spans for the United States, the pioneer of natural gas use, were shorter but still considerable: 53 years to reach 5 percent, another 31 years to get to 25 percent.

Displacing even just a third of today's fossil fuel consumption by renewable energy conversions will be an immensely challenging task; how far it has to go is attested by the most recent shares claimed by modern biofuels and by wind and photovoltaic electricity generation. In 2010 ethanol and biodiesel supplied only about 0.5 percent of the world's primary energy, wind generated about 2 percent of global electricity and photovoltaics (PV) produced less than 0.05 percent. Contrast this with assorted mandated or wished-for targets: 18 percent of Germany's total energy and 35 percent of electricity from renewable flows by 2020, 10 percent of U.S. electricity from PV by 2025 and 30 percent from wind by 2030 and 15 percent, perhaps even 20 percent, of China's energy from renewables by 2020.

Unit sizes of new converters will not make the transition any easier. Ratings of 500–800 megawatts (MW) are the norm for coal-fired turbogenerators and large gas turbines have capacities of 200–300 MW, whereas typical ratings of large wind turbines are two orders of magnitude smaller, between 2 and 4 MW, and the world's largest PV plant needed more than a million panels for its 80 MW of peak capacity. Moreover, differences in capacity factors will always remain large. In 2009 the load factor averaged 74 percent for U.S. coal-fired stations and the nuclear ones reached 92 percent, whereas wind turbines managed only about 25 percent—and in the European Union their mean load factor was less than 21 percent between 2003 and 2007, while the largest PV plant in sunny Spain has an annual capacity factor of only 16 percent.

As I write this a pronounced high pressure cell brings deep freeze, and calm lasting for days, to the usually windy heart of North America: If Manitoba or North Dakota relied heavily on wind generation (fortunately, Manitoba gets all electricity from flowing water and exports it south), either would need many days of large imports—yet the mid-continent has no high-capacity east-west transmission lines. Rising shares of both wind and PV generation will thus need considerable construction of new long-distance high-voltage lines, both to connect the windiest and the sunniest places to major consumption centers and also to assure uninterrupted supply when relying on only partially predictable energy flows. As the distances involved are on truly continental scales—be they from the windy Great Plains to the East Coast or, as the European plans call for, from the reliably sunny Sahara to cloudy Germany (Desertec plan)—those expensive new supergrids cannot be completed in a matter of years. And the people who fantasize about imminent benefits of new smart grids should remember that the 2009 report card on the American infrastructure gives the existing U.S. grid a near failing grade of D+.

And no substantial contribution can be expected from the only well-tested non-fossil electricity generation technique that has achieved significant market penetration: Nuclear fission now generates about 13 percent of global electricity, with national shares at 75 percent in France and about 20 percent in the United States. Nuclear engineers have been searching for superior (efficient, safe and inexpensive) reactor designs ever since it became clear that the first generation of reactors was not the best choice for the second, larger, wave of nuclear expansion. Alvin Weinberg published a paper on inherently safe reactors of the second nuclear era already in 1984, at the time of his death (in 2003) Edward Teller worked on a design of a thorium-fueled underground power plant, and Lowell Wood argues the benefits of his traveling-wave breeder reactor fueled with depleted uranium whose huge U.S. stockpile now amounts to about 700,000 tonnes.

But since 2005, construction began annually on only about a dozen new reactors worldwide, most of them in China where nuclear generation supplies only about 2 percent of all electricity, and in early 2011 there were no signs of any western nuclear renaissance. Except for the completion of the Tennessee Valley Authority's Watts Bar Unit 2 (abandoned in 1988, scheduled to go on line in 2012), there was no construction underway in the United States, and the completion and cost overruns of Europe's supposed new showcase units, Finnish Olkiluoto and French Flamanville, were resembling the U.S. nuclear industry horror stories of the 1980s. Then, in March 2011, an earthquake and tsunami struck Japan, leading to Fukushima's loss of coolant, destruction of reactor buildings in explosions and radiation leaks; regardless of the eventual outcome of this catastrophe, these events will cast a long suppressing shadow on the future of nuclear electricity.

Technical Fixes to the Rescue?

New energy conversions are thus highly unlikely to reduce CO_2 emissions fast enough to prevent the rise of atmospheric concentrations above 450 parts per million (ppm). (They were nearly 390 ppm by the end of 2010). This realization has led to enthusiastic

exploration of many possibilities available for carbon capture and sequestration—and to claims that would guarantee, even if they were only half true, futures free of any carbon worries. For example, a soil scientist claims that by 2100 biochar sequestration (essentially converting the world's crop residues, mainly cereal straws, into charcoal incorporated into soils) could store more carbon than the world emits from the combustion of all fossil fuels.

Most of these suggestions have been in the realm of theoretical musings: Notable examples include hiding CO_2 within and below the basalt layers of India's Deccan (no matter that those rocks are already much weathered and fractured), or in permeable undersea basalts of the Juan de Fuca tectonic plate off Seattle (but first we would have to pipe the emissions from Pennsylvania, Ohio and Tennessee coal-fired power plants to the Pacific Northwest)—or using exposed peridotites in the Omani desert to absorb CO_2 by accelerated carbonization (just imagine all those CO_2-laden megatankers from China and Europe converging on Oman with their refrigerated cargo).

One of these unorthodox ideas has been actually tried on a small scale. During the (so far) largest experiment with iron enrichment of the surface ocean (intended to stimulate phytoplankton growth and sequester carbon in the cells sinking to the abyss) an Indo-German expedition fertilized of 300 square kilometers of the southwestern Atlantic in March and April 2009—but the resulting phytoplankton bloom was devoured by amphipods (tiny shrimp-like zooplankton). That is why the best chances for CCS are in a combination of well-established engineering practices: Scrubbing CO_2 with aqueous amine has been done commercially since the 1930s, piping the gas and using it in enhanced oil recovery is done routinely in many U.S. oilfields, and a pipeline construction effort matching the extension of U.S. natural gas pipelines during the 1960s or 1970s could put in place plenty of links between large stationary CO_2 sources and the best sedimentary formations used to sequester the gas.

But the scale of the effort needed for any substantial reduction of emissions, its safety considerations, public acceptance of permanent underground storage that might leak a gas toxic in high concentrations, and capital and operation costs of the continuous removal and burial of billions of tonnes of compressed gas combine to guarantee very slow progress. In order to explain the extent of the requisite effort I have been using a revealing comparison. Let us assume that we commit initially to sequestering just 20 percent of all CO_2 emitted from fossil fuel combustion in 2010, or about a third of all releases from large stationary sources. After compressing the gas to a density similar to that of crude oil (800 kilograms per cubic meter) it would occupy about 8 billion cubic meters—meanwhile, global crude oil extraction in 2010 amounted to about 4 billion tonnes or (with average density of 850 kilograms per cubic meter) roughly 4.7 billion cubic meters.

This means that in order to sequester just a fifth of current CO_2 emissions we would have to create an entirely new worldwide absorption-gathering-compression-transportation- storage industry whose annual throughput would have to be about 70 percent larger than the annual volume now handled by the global crude oil industry whose immense infrastructure of wells, pipelines, compressor stations and storages took generations to build. Technically possible—but not within a timeframe that would prevent CO_2 from rising above 450 ppm. And remember not only that this would contain just 20 percent of today's CO_2 emissions but also this crucial difference: The oil industry has invested in its enormous infrastructure in order to make a profit, to sell its product on an energy-hungry market (at around \$100 per barrel and 7.2 barrels per tonne that comes to about \$700 per tonne)—but (one way or another) the taxpayers of rich countries would have to pay for huge capital costs and significant operating burdens of any massive CCS.

And if CCS will not scale up fast enough or it will be too expensive we are now offered the ultimate counter-weapon by resorting to geoengineering schemes. One would assume that a favorite intervention—a deliberate and prolonged (decades? centuries?) dispensation of millions of tonnes of sulfur gases into the upper atmosphere in order to create temperature-reducing aerosols—would raise many concerns at any time, but I would add just one obvious question: How would the Muslim radicals view the fleets of American stratotankers constantly spraying sulfuric droplets on their lands and on their mosques?

These are uncertain times, economically, politically and socially. The need for new departures seems obvious, but effective actions have failed to keep pace with the urgency of needed changes—particularly so in affluent democracies of North America, Europe and Japan as they contemplate their overdrawn accounts, faltering economies, aging populations and ebbing global influence. In this sense the search for new energy modalities is part of a much broader change whose outcome will determine the fortunes of the world's leading economies and of the entire global civilization for generations to come. None of us can foresee the eventual contours of new energy arrangements—but could the world's richest countries go wrong by striving for moderation of their energy use?

Bibliography

British Petroleum. 2010. BP Statistical Review of World Energy. http://www.bp.com/liveassets/bp_internet/globalbp/globalbp_uk_english/reports_and_publications/statistical_energy_review_2008/STAGING/local_assets/2010_downloads/statistical_review_of_world_energy_full_report_2010.pdf

International Energy Agency. 2010. *Key World Energy Statistics*.

http://www.iea.org/textbase/nppdf/free/2010/key_stats_2010.pdf

Jackson, R. B., and J. Salzman. 2010. Pursuing geoengineering for atmospheric restoration. *Issues in Science and Technology* 26:67-76.

Metz, B., et al., eds. 2005. *Carbon Dioxide Capture and Storage*. Cambridge: Cambridge University Press.

Sivak, M., and O. Tsimhoni. 2009. Fuel efficiency of vehicles on US roads: 1923-2006. *Energy Policy* 37:3168-3170.

Smil, V. 2008. *Energy in Nature and Society: General Energetics of Complex Systems*. Cambridge, MA: The MIT Press.

Smil, V. 2009. U.S. energy policy: The need for radical departures. *Issues in Science and Technology* 25:47-50.

Smil, V. 2010. *Energy Transitions: History, Requirements, Prospects*. Santa Barbara, CA: Praeger.

Smil, V. 2010. *Energy Myths and Realities: Bringing Science to the Energy Policy Debate*. Washington, DC: American Enterprise Institute.

U.S. Energy Information Administration. 2010. International Energy Outlook 2010. http://www.eia.doe.gov/oiaf/ieo/pdf/0484(2010).pdf

For relevant Web links, consult this issue of *American Scientist Online:*

http://www.americanscientist.org/issues/id.90/past.aspx

Marking Loons, Making Progress

Striking discoveries about the social behavior and communication of common loons are revealed by a low-tech approach: individual marking of study animals

Walter Piper, Jay Mager and Charles Walcott

Anyone who has spent time on lakes in the northern United States or Canada is familiar with the common loon. Its haunting calls are emblematic of wild places. They captivate the listener. The ringing tremolo, for example, suggests maniacal laughter; the wail recalls the distant cry of a person in distress. Despite their humanlike qualities, though, the vocalizations of loons are acoustically complex and not easily deciphered by humans.

Even when not vocalizing, loons are charismatic creatures. They are striking in appearance: clad in black with lines of bright white ovals on their wings, intricate white badges on the sides of their long necks and delicate white "chinstraps." Their burning orange eyes seem to miss no detail. They are most often encountered foraging, making dives of a minute or more in duration to pursue fish and invertebrates underwater, propelled by their powerful legs. After watching the speed and agility of a loon passing under one's canoe, one can only pity a sunfish! Naturally enough, loons are loved and admired by humans who share their northern lakes.

The common loon is the most abundant and widespread of five species of loons that are distributed throughout the Northern Hemisphere, breeding across Canada and the northern United States and wintering along the Pacific and Atlantic coasts. Loons are socially monogamous: Pairs defend territories on freshwater lakes and build nests on islands or shorelines, where the female of the pair lays two eggs. Most clutches fall prey to raccoons, skunks and other predators during the four weeks of joint incubation by male and female. The chicks that survive to hatching are semiprecocial—downy, eyes open, and able to leave the nest to swim and dive soon after emerging. Chicks are warm brown in color, which allows them to hide effectively among rocks and logs near shore while parents are engaged in foraging and social interactions with other loons. Parents feed and attend to chicks constantly for 11 weeks, after which the chicks can feed themselves.

Most of our knowledge of common loon biology was expertly summarized in Judith McIntyre's 1989 book, *The Common Loon: Spirit of Northern Lakes.* When it was published, the book was as remarkable for what it could not say as for what it could. Missing or rudimentary were an understanding of the mating system, territory acquisition by young adults, territory defense by established breeders, and, tantalizingly, the functions of loon vocalizations.

The principal reason for our ignorance about loon behavior was simple: Observers could not identify and distinguish the individuals they studied. One might think the intricate plumage patterns of adult loons would make identification easy, and loons do display distinctions in fine plumage details, such as the number of lines and branches in neck badges, but these features can only be reliably distinguished with the bird in hand. Instead of serving as dependable ID badges, loon markings can create the illusion in observers' minds that the males and females they observe each year are the same as those seen the year before, when in fact that is not certain. Complex behaviors such as calls and social interactions may seem to dependably distinguish a familiar bird to attentive observers—but even apparently peculiar behavior may be stereotyped and thus deceptive as an identifying feature.

Band Practice

The assumption of many observers that loon pairs mate for life could be put to the test only after 1992, when David Evers and his team refined the nocturnal spotlighting technique and began to capture and mark hundreds of adults each year with colored leg bands. In this method, animals are located with a spotlight, then researchers creep up on them with an electric motor boat. Adults can be confused and frozen in place by an imitation of the chick distress call (a hoarse, plaintive whistle). Chicks can be immobilized by the "hoot" contact call that adults give each other and chicks. Both can then be scooped up with nets.

Evers and his colleagues showed not only that one could capture loons, despite their large size (4–5 kilograms) and tendency to dive when threatened, but that one could do so with sufficient frequency and predictability to maintain a marked study population. The time for systematic investigation of loon behavior had come.

By cobbling together breeding pairs banded by Evers's team with some that we banded ourselves using his technique, we began to examine loon territorial behavior in 1993 in Oneida County, Wisconsin. As an investment in the future, we banded not only adult loons but also chicks of five weeks of age or older, whose legs had grown large enough to hold the plastic bands safely. Our banding and observation occurred within a cluster of about 100 glacial lakes ranging from 4 to over 500 hectares in size comprising a roughly circular area 50 kilometers in diameter.

Walter Piper is professor of earth and environmental sciences at Chapman University. Charles Walcott is professor emeritus of neurobiology and behavior at Cornell University. Jay Mager is an assistant professor of biological and allied health sciences at Ohio Northern University. Address for Piper: Chapman University, Orange, CA 92866. Email: wpiper@chapman.edu

Figure 1. Some study organisms are more charismatic than others. The common loon, *Gavia immer,* is talkative, athletic, handsome and until recently somewhat mysterious. The development of an efficient capture and banding technique has allowed long-term study over the loon's life span, which can pass two decades. After 18 years of study on banded populations, there are still questions remaining about the progression from chick to battler to homesteader to ancient, but there has also been a harvest of answers. The authors review current knowledge of the domestic arrangements, real estate transactions and conversations of these captivating animals. Photograph courtesy of Carol Henderson.

Most lakes in this area are surrounded by summer cottages and have high boating activity. Thus, although scenic beauty is diminished on human-impacted lakes, loon pairs that live and breed on them are tame enough that we can approach within 5 meters to identify individuals by their leg bands and record their social behavior.

Observation of marked loons on their territory allowed us to put to the test the popular conception that loons mate for life. Indeed, we soon learned that male or female pair members—always one or the other, not both—routinely vanished from their territories, leaving, for example, a marked male and unmarked female where there had previously been a marked pair. A few such disappearances would not have been surprising. After all, loons are subject to disease, injury and predation, just as other animals are, and we might expect that a breeder that died would be replaced by a new breeder. In most cases, however, we located the missing pair member on a vacant lake immediately adjacent to its original territory. The behavior of these birds was sometimes more subdued than it had been—more resting, less foraging and social activity—and in some cases obvious lacerations or defeathered patches were visible on their heads or necks. In time, we accumulated enough such observations that we began to understand that both males and females were in a constant struggle to protect their territories from intruders, which sometimes were able to evict a resident and seize the territory for themselves.

Unwelcome Mat

The regular eviction of owners from their territories caused us to turn our attention to a segment of the loon population that is obvious even in the absence of marked birds: floaters. Floaters—male and female adults that lack territories but otherwise resemble territory owners—are abundant within the population. Though initially we had no idea about the age or origin of these floaters, since they were unbanded, their visits to the lakes defended by territorial pairs are routine, averaging 2–5 visits per day throughout the breeding season (April to August). The intrusions we observe most often are highly stereotyped. They begin when a floater lands in the lake. The original pair becomes highly alert, their heads held almost comically high above the water as they scour the surface for the foreigner. The pair then proceed in tandem toward the intruder. Almost invariably, the three loons converge, engaging in a series of social behaviors—head bows, circle dances and splash dives. The intruder leaves within 30 minutes. Although the lack of marked floaters made our conclusions speculative, it seemed that a typical floater spent much of its time systematically intruding into established territories, staying from 5 to 30 minutes and then departing.

Figure 2. Loons build their nests close to the water. If threatened they can simply slide off the nest and disappear below the surface of the water. The incubating loon above is lowering its head because of the presence of the photographer, a common loon behavior when a human approaches its nest. This behavior reduces the visibility of the loon to a potential predator on the lake. Photograph courtesy of Dan Salisbury.

What might floaters gain from such territorial visits? The energetic investment required to fly between many lakes, find territorial pairs and interact with them seemed too great for birds to engage in without reward, yet most intrusions were seemingly peaceful affairs, not knock-down-drag-out contests for territorial ownership. Although we seldom observed social interaction between male intruders and female pair members, it seemed plausible that male intruders were not usually bent on usurping the territory, but instead were interested in mating with the resident female and thus siring young without rearing them. The avian literature abounds with such examples of "extrapair parentage," especially in songbirds, and many scientists have found it despite little or no behavioral evidence of "extrapair matings." Actually, one might expect that a female pair member who mated outside of her pair bond—perhaps as a means to ensure fertile eggs or to improve the genes of her offspring—would do so furtively, so as not to cause her mate to suspend parental care. To explore the possibility of extrapair parentage, we took blood samples from males, females and their one or two chicks and used DNA fingerprinting to analyze parentage. The results were clear: not a single case of extrapair parentage in 58 young from 47 families.

While we might have hoped for more colorful findings, we had all but eliminated the possibility that floaters were males systematically intruding into territories to attempt to copulate with the breeding females.

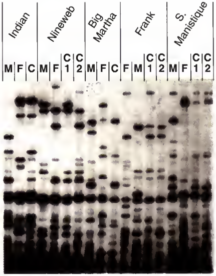


Figure 3. Multilocus DNA fingerprints of five loon families, including males ("M"), females ("F") and chicks ("C", "C1", "C2"). Fragments are passed from parents to offspring in Mendelian fashion. Thus the overall similarity of parents' fingerprints to their offspring's reveals the statistically confirmed conclusion of genetic monogamy.

Hostile Takeover

If intruders were not males looking to mate secretly with the territorial females, what were they doing? The answer was not obvious. Still, by analyzing rates of intrusion into each territory and tracking the tendency of each territory to produce chicks, we discovered a link. In years following chick production by a territorial pair, the rate of intrusions increases by 60 percent. Moreover, territorial takeovers—by males and females both—are also significantly more common in years following chick production. In short, intruders appear to target recently successful territories for takeover and often succeed in taking over those territories. Why are recently successful territories favored? Because reproductive success is highly autocorrelated in loon territories. That is, a territory that produces chicks in year 1 is very likely to produce chicks in year 2, either because chick production indicates a high-quality territory or because conditions favorable for producing chicks (such as a low density of raccoons) tend to last for several years. So

an intruder that succeeds in usurping a successful territory one year is likely to be rewarded with a favorable environment for breeding the subsequent year.

Although simple observation of loon pairs and temporal patterns of territorial intrusion gave us great insights into territory acquisition, we have only now started to achieve a fuller picture of the process. Birds that we marked as chicks have returned to the study area as adults (having acquired their adult plumage after a year or two residing as juveniles in salt-water areas) and begun to usurp territories. Male floaters that are 3 to 4 years old, we have learned, almost always acquire a territory by "founding"—settling in a vacant lake or part of a large lake and pairing with an available female. At age 5 or 6, males suddenly begin to battle for established territories. Finally, old males—those of at least 15 years—behave like very young males: They found territories, eschewing open combat. Females, on the other hand, show no relationship between age and mode of territory acquisition, being just as likely at all ages to acquire territories by settling without conflict in a vacant territory or by wresting ownership of a territory by force.

Capture and recapture of marked individuals has allowed us to document a change in the loons' body condition as they age that parallels changes in territory acquisition. Males have relatively low body mass when 4 and 5 years old, which increases steadily from 6 to 10 and then declines late in life. Female body mass fluctuates very little with age. The correlation between mode of gaining a territory and male body mass suggests that males avoid territorial confrontations when their body condition is marginal and take advantage of good body condition during young adulthood to seize territories by force. Males have good reason to avoid contests for territory ownership when not in prime condition: About a third of male takeovers (but not female ones) result in the death of a combatant, nearly always the original resident.

Figure 4. Loons are sturdy, dense birds, which facilitates plunging dives for fish but makes taking flight a challenge. They need a substantial water runway to do so. The loon's legs are set far to the rear, which is a fine configuration for swimming but nearly useless for walking on land. A loon that mistakes a wet road for a lake is doomed unless it can convey itself to a large enough body of water. Photograph courtesy of Dan Salisbury.

At All Costs

The observation of lethal combat was, in fact, a great surprise. It seems to be quite rare in animals—although this conclusion is tentative, as most field studies have limited capacity to observe lethal contests and recover corpses of slain individuals. Most fatal fighting has been detected in short-lived species, like fig wasps and spiders—groups wherein an individual might get only one opportunity to reproduce and should be expected to battle mightily for that opportunity. Loons do not fit this model, as they often live upwards of twenty years. Why would an animal with a long life expectancy ever get caught in a battle that might cost his life? Can a territory be worth so much that a male should risk death to hold onto it?

One hypothesis that might help explain fatal fighting in male loons holds that the value of a territory to its owner might increase greatly over time, justifying vigorous territorial defense. The key to understanding a male's stake in his territory lies in nesting behavior. When a pair nests and succeeds, they tend to reuse the successful site, which makes sense, as its success likely indicates a location safe from nest predators. When a pair fails, they usually move the nest. This intuitive "win–stay, lose–switch"

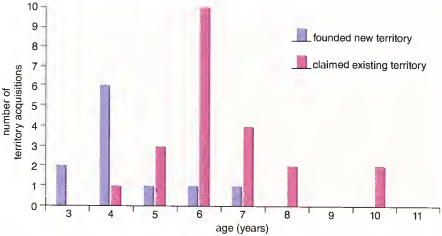


Figure 5. Males show stark changes in their mode of territory acquisition with increasing age. Males 3 to 4 years old found new territories in vacant lakes or parts of lakes; males of 5 to 8 years almost always claim an occupied territory, which requires them to evict the current resident.

Figure 6. This fight between two male loons ended with a 16-minute chase across the surface of the water. About 30 percent of male battles end with a fatality, always the resident male. Analysis of this intriguing finding suggests that loons make sophisticated, high-stakes investments in defense of nesting sites and acquired familiarity with local territories. At right, a 3-year-old banded intruder flees after a brief visit to a defended territory. Photographs by Dan Salisbury.

rule is widespread in vertebrates and has been shown to improve nesting success in loons and many other species. Since we had marked breeding pairs, we could ask, "Which pair member is using the rule?" We answered this question by looking at use of the nesting rule under different scenarios.

When the male from the previous year returned but the female had been replaced by a new bird, the pair continued to use the rule—as did pairs in which both male and female from the previous season had returned. On the other hand, new male/old female pairs ceased reusing successful locations, and, indeed, selected nest sites with no more success than new male/new female pairs. Moreover, pairs containing a male with past breeding success on the territory were 41 percent more likely to produce chicks than were pairs with an inexperienced male.

From a male's perspective, then, nesting on a familiar territory is a good deal better than nesting on an unfamiliar one, because he knows where to nest and is far more likely to produce young there. Hence, one can see why a male who knows where to nest on his territory might fight furiously to keep it. A territorial male that accepts defeat by an intruder vying for his territory accepts a sharp decrease in reproductive success.

The "familiarity" hypothesis for fatal fighting among males generates one clear prediction that we can examine immediately: The males that perish during territorial contests should usually be resident males—those with a large stake in staying put. Intruders that challenge them should be less likely to be truly dogged in a territorial contest, as they do not know the territory and will have to learn by trial and error how to use it in the event that their eviction attempt succeeds. We quickly confirmed this prediction, as nearly all males found dead after territorial battles were territorial owners. Combat is a mutual endeavor, like a dance, and yet the owners were the only combatants that fought unto death. The record of numerous combats shows no such commitment from invaders. Combining this finding with our recent findings with respect to male territory acquisition, we can infer that fatal battles are those that pit an experienced owner, who might have declined a bit in body condition from his prime, against a young, fit, 5- to 8-year-old floater.

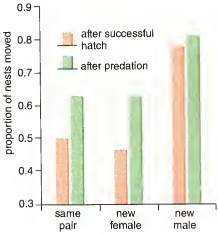


Figure 7. The graph compares movement of nest location between consecutive nesting attempts by loon pairs of different compositions. If both male and female returned from the previous year ("same pair"), pairs tended to position the nest according to the success of their last attempt, reusing successful sites and abandoning ones where eggs had been predated. Turnover of the female pair member ("new female") did not alter the use of this nesting strategy. But turnover of the territorial male from the previous year resulted in nests being relocated, regardless of success.

As satisfying as the familiarity hypothesis seems, it has a glitch. A 41 percent bonus in reproductive success for the resident male is not trivial, but it seems an insufficient payoff to make a male risk his life. After all, even a male that accepted eviction from his territory and was forced to move to a new one would likely recover his lost reproductive success within 3 to 5 years. Why risk death?

Figure 9. The soundtrack of the wild in many northern territories, the yodeling of the male loon is an announcement of territoriality consisting of an introductory phrase followed by repeat syllables. Male loons have a consistent yodel year after year, unique to themselves, but if they change territories, they often change their yodel. Photograph courtesy of Nathan Banfield.

The "desperado" hypothesis might help explain fatal fighting. Perhaps the decline we have detected in older males is rapid enough that there might be some greybeards out there with just a year or two left to live. If so, and if old males can *detect* their loss of condition, they might well become "desperados"—individuals with so little time left that it becomes worthwhile risking their lives for another year or two on a familiar territory, instead of accepting displacement to an unknown territory for a year or two of marginal reproduction. While the hypothesis seems sound, we are still collecting enough data on very old males of known age to test the desperado hypothesis robustly. In this case, 18 years of data is not enough!

Caller ID

The dramatic changes that occur during the lifetime of a territorial male raise the possibility that the male-only territorial call, the yodel, might also change as a male ages. A ringing advertisement of territorial behavior, the yodel consists of an introductory phrase that rises in pitch followed by one or more "repeat syllables," which comprise two brief, frequency-modulated notes. Investigation within and between breeding seasons has shown that each male's yodels are unique in terms of timing, frequency and frequency modulation. In effect, the yodel of a male is his "voice," recognizable to humans who know it and probably also to other loons. The variability and acoustic complexity of the yodel suggests that it might contain information about the signaler in addition to his identity.

At first glance, it seems quite likely that the yodel conveys information about a male's ability to defend a territory. Almost all yodels occur when a male is in a tense interaction with an intruder that has landed on its territory or when an intruder is flying overhead. So yodels appear to be "aimed" at specific territorial opponents. If so, we should expect that a yodeler wishes to communicate something about himself or his motivation to defend his territory—perhaps in an effort to save himself the trouble and energetic cost of a lengthy confrontation or battle. One crucial bit of information that a male might want to communicate is his body size or condition. We might hypothesize that a large or fit male, who would be a formidable opponent, would want to communicate that, if possible, through his yodel.

In fact yodels do betray information about the fighting ability of the yodeler. The information is encoded in the dominant frequency of the yodel—the frequency of the yodel that is of greatest amplitude, or loudest. However, dominant frequency is closely correlated not with body size but with body mass. Thus, it seems that heavy males are potentially signaling their body condition through their lower frequency yodel. We also found a correlation between change in body mass and dominant frequency from one year to the next. That is, males whose yodels rise in pitch from one year to the next have lost body mass over that period and those whose yodels become deeper have gained mass. A loon listening to a male's yodel, therefore, could instantly get information about its condition, and, if it recognized the same male yodel from the year before, could learn whether the male was gaining or losing condition.

It is one thing for humans to use acoustic analysis software to learn that features of loon calls contain information about the loons themselves and quite another to show that loons listening to the signal actually perceive and act on this information. To see whether receivers might use information about body condition within loon calls, we played three versions of a stranger's yodel to territory holders: 1) the original yodel, 2) the original yodel decreased in frequency by 200 hertz (a change that left the altered call within the normal range of loon yodels), and 3) the original yodel increased by 200 hertz (which also remained within the natural frequency

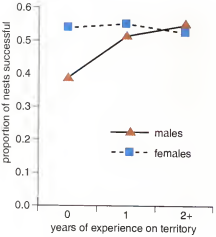


Figure 8. The nesting success of males and females is compared according to their length of experience on a territory. As a consequence of using the "win–stay, lose–switch" rule, males improve their breeding success markedly between their first and later years. Females, which do not control nest placement, show no such improvement.

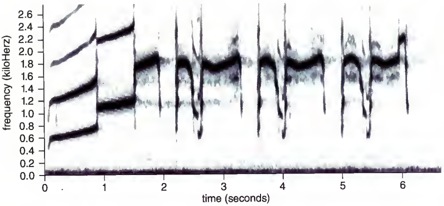


Figure 10. This sonogram of a male loon's yodel shows the introduction followed by three repeat syllables at the end. Measuring the frequencies of two notes and the duration of the introduction and space between the notes allows us to tell different loons apart by their yodel. Analysis of many yodels indicates that the loon is announcing his identity, fitness and willingness to fight.

range). We found that territorial males responded more quickly and aggressively towards yodels that were shifted to lower frequencies. This result suggests that loons view low-frequency yodels as more threatening. They really do pay attention to frequency! It remains to be seen, however, whether loons listening to yodels have the capacity to remember and learn from differences in yodels compared across years. If so, they might use this ability to target territories of declining owners for takeover attempts.

There is one puzzling aspect to the finding of "honest signaling" of body size, as animal behaviorists call it. Why would a small male ever yodel, if in so doing he was revealing his small size to all listeners? This practice would seem masochistic—equivalent to asking another male to try to evict you from your territory. Yet small males yodel about as often as do large males. There must be some compensating benefit to yodeling that offsets the disadvantage of telling competitors that you are small.

Indeed, there is more information contained in the yodel than just identity and condition. The mere fact that a male is yodeling reveals that he is territorial, of course, but we wanted to determine whether the number of repeat syllables in a male's yodels communicated something about his aggressiveness. Repeat syllables vary greatly in number from yodel to yodel within and between males, so it seemed plausible that this flexible aspect of the call was being used for this purpose. Indeed, we have learned that males produce more repeat syllables in their yodels when in close encounters with intruders than when intruders were flying over or far away on the water. We followed up this promising piece of field data with an experiment that simulated the intrusion of a male floater into the territory defended by a breeding pair. As expected, territorial males responded more quickly and with more tremolos and yodels of their own to foreign yodels manipulated to have four or seven repeat syllables than to those altered to have only one repeat syllable. Thus, the number of repeat syllables in the yodel is taken by other loons as a signal of high aggressive motivation on the part of the yodeler. So a small male must yodel in order to signal his territorial status and aggressiveness, even though his yodel betrays his small size in the process.

We performed a second experiment on aggressive motivation in which we actually altered the quality of a male's territory. Although this might sound daunting, one can improve territory quality simply by constructing a 1 × 1 meter floating platform from cedar logs, adding buoyant material to improve flotation and anchoring this "nesting raft" 20 meters or more from shore. We have shown that nesting platforms increase hatching success by 69 percent by deterring raccoons and other egg predators. Males in territories with added platforms emitted longer yodels compared to their own calls in the years before and after platform deployment and compared to those of control males whose territories had no platforms. These results suggest that males crank up their territory defense when their territory becomes valuable, providing further evidence of how males use yodels to communicate aggressive motivation.

Our use of marked birds allowed us to track males that moved from one territory to another, and it exposed an unexpected and baffling aspect of their yodels: The yodel changes when a male moves from one territory to another. Specifically, yodels change in the timing and peak frequency of the introductory note and the length of the

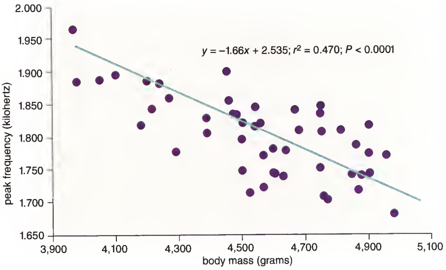


Figure 11. The peak frequencies for yodels of males with low body mass are much higher than those of heavy males, raising the question of why a smaller male would engage in "honest signaling" that would seem to advertise small size—no small matter between competing territorial animals. Possibly the smaller animal is also signaling his aggressiveness in the duel for territory, also no small matter.

delay between the introductory note and first repeat syllable. What does a male gain from shifting his yodel in this way when he claims a new territory? Is he hiding his identity? If so, this might help him evade young males that had been targeting his previous territory for takeover attempts. We would predict, in this case, that the new yodels would be unrelated to their original form, not just modulated in some systematic way that could be decoded. But the changes in males yodels *are* systematic. Shifts in acoustic parameters of the yodels almost always make the male's yodel differ as much as possible from that of the male that resided on his new territory before him. This odd finding supports the hypothesis that males shift their yodels to proclaim that they are new to the territory and should not be confused with the previous male. Such a strategy might make sense. The old male, after all, was humbled and possibly killed, suggesting that he might have been in poor condition and subject to many eviction attempts by floaters. Complete testing of these hypotheses will occur when we detect two or three times as many territory and yodel shifts, at which point we will be able to link the cause of the territorial turnover with the change in the yodel and search also for an impact of the change in ownership on the rate of territorial intrusion.

We have made great strides in understanding how loons gain and defend breeding territories, following systematic capture and marking of our study animals. Marking of individuals has revealed that territorial takeover is common; that among males, only those in prime condition achieve takeovers; and that male loons control where the nest is placed in a breeding attempt, thus acquiring unique familiarity with their territory and increasing the value of the territory to the male. Our vocal analysis has been similarly dependent on recording the calls of marked males. Field observations and playback experiments have shown that yodels betray the yodeler's size and aggressive motivation, and have revealed the curious alteration of yodels by males that have just taken possession of a new territory. Much about loon behavior remains unknown, but systematic marking and study of the species has given us the power to unravel many of the mysteries of this most engaging animal.

Bibliography

Evers, D. C. 1993. Population ecology of the common loon at the Seney National Wildlife Refuge, Michigan: Results from the first color-marked breeding population. In Morse, L, S. Stockwell and M. Pokras (eds). *The Loon and Its Ecosystem.* Concord, NH: U.S. Fish and Wildlife Service, 202–212.

Grafen, A. 1987. The logic of divisively asymmetric contests: respect for ownership and the desperado effect. *Animal Behavior* 35:462–467.

Mager, J. N., C. Walcott and W. H. Piper. 2007. Male common loons, *Gavia immer*, communicate body mass and condition through dominant frequencies of territorial yodels. *Animal Behavior* 73:683–690.

Mager, J. N., C. Walcott and W. H. Piper. 2008. Nest platforms increase aggressive behavior in common loons. *Naturwissenschafte* 95:141–147.

McIntyre, J. W. 1988. *The Common Loon: Spirit of Northern Lakes.* Minneapolis: University of Minnesota Press.

Piper, W. H., D. C. Evers, M. W. Meyer , K. B. Tischler, J. D. Kaplan and R. C. Fleischer. 1997. Genetic monogamy in the common loon (*Gavia immer*). *Behavioral Ecology and Sociobiology* 41:25–31.

Piper, W. H., M. W. Meyer, M. Klich, K. B. Tischler and A. Dolsen. 2002. Floating platforms increase reproductive success of common loons. *Biological Conservation* 104:199–203.

Piper, W. H., K. B. Tischler and M. Klich. 2000. Territory acquisition in loons: the importance of take-over. *Animal Behavior* 59:385–394.

Piper, W. H., C. Walcott, J. N. Mager and F. J. Spilker. 2008. Fatal battles in common loons: A preliminary analysis. *Animal Behavior* 75:1109–1115.

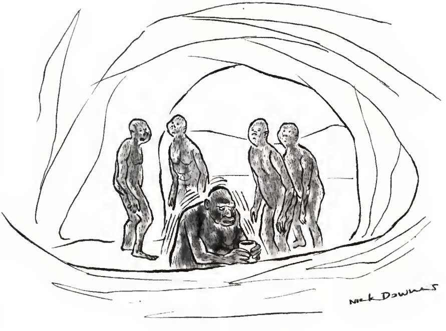
Piper, W. H., C. Walcott, J. N. Mager and F. J. Spilker. 2008. Nestsite selection by male loons leads to sex-biased site familiarity. *Journal of Animal Ecology* 77:205–210.

Reece, S.E., T. M. Innocent and S. A. West. 2007. Lethal male-male combat in the parasitoid *Melittobia acasta*: Are size and competitive environment important? *Animal Behavior* 74:1163–1169.

Walcott, C., J. N. Mager and W. Piper. 2006. Changing territories, changing tunes: male loons, *Gavia immer*, change their vocalizations when they change territories. *Animal Behavior* 71:673–683.

For relevant Web links, consult this issue of *American Scientist Online:*

http://www.americanscientist.org/Issues/id.90/past.aspx



"Him we call 'Java-Man.'"

Pliocene Climate Lessons

An ongoing reconstruction of a warmer Earth 3 million years ago helps test climate-change forecasts

Marci Robinson

Over the course of geologic history, global temperatures have increased and decreased in response to a multitude of forces. Changes in the positions of continents, which can modify circulation in ocean basins and affect the global distribution of heat, have been a major cause of global climate change. Variations in Earth's orbital configurations, which guide the glacial-interglacial cycles, have been important. So have changes in the atmospheric concentration of carbon dioxide, a greenhouse gas that traps heat in our atmosphere. Earth's present warming trend is happening too quickly to be related to the first two factors. So most scientists attribute the change to the rapid increase in greenhouse gases—primarily from fossil-fuel combustion—accumulating in our atmosphere.

What will this mean for our future? Answering that question has fallen in large part to the Intergovernmental Panel on Climate Change (IPCC), the international body that synthesizes volumes of climate research into a single report every five to six years. In 2007, the IPCC Fourth Assessment Report predicted a mean annual global temperature increase of between 1.1 degrees and 6.4 degrees Celsius by the end of the 21st century. These conclusions were arrived at using general circulation model, or GCM, experiments. GCMs are numerical models of the planet that simulate physical processes in the ocean, atmosphere, cryosphere and at the surface.

Such a wide range of predicted temperature change, of course, produces a wide range of climate-effect forecasts. A climate that is on average 1 degree warmer than today is likely to be very different than a climate that is 6 degrees warmer. That could represent the difference between 18 centimeters and 59 centimeters of sea-level rise, for example, in the next 90 years. How much faith should be placed in different model results? How do we quantify their uncertainty?

One way to seek a more accurate view of the future is to look closely at the past, to a time when Earth's temperatures matched what may occur by the end of this century. Temperature records, dating only from 1850, don't carry us back far enough. Nor can ice cores, which capture atmospheric carbon dioxide levels dating back only 800,000 years. But studies based on chemical ratios preserved in tiny marine fossils, along with other proxies, are filling in some blanks. The results are becoming a useful "ground truth" for climatologists attempting to model our future climate. At the same time, the models are helping paleoclimate researchers such as myself improve our understanding of the past.

Pliocene Proxy

The most powerful proxy tool available to help scientists identify past warm periods is the marine oxygen isotope record, which is based on the ratio of two oxygen isotopes—oxygen-16 and oxygen-18—preserved in the shells of microfossils. During evaporation, the lighter isotope—^{16}O—is preferentially removed from the water, leaving the remaining sea water enriched with ^{18}O. The evaporated water contributes to the snow that builds ice caps. During cold periods, ice caps grow, trapping the lighter isotope, and sea level falls, which leaves seawater enriched in the heavier isotope.

Marine organisms preserve the oceans' isotopic signatures when they precipitate calcium-carbonate shells. Shells formed during glacial periods have a higher proportion of heavy oxygen than shells formed during interglacial periods. Because changes in the isotopic signature are synchronous across all oceans, records from different parts of the world can be correlated. The efforts of many scientists analyzing millions of shells spanning millions of years of geologic time have yielded an oxygen-isotope record that roughly estimates global changes in temperature and sea level.

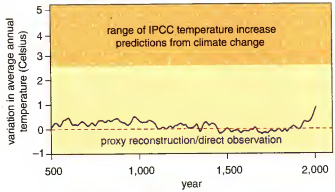
According to this record, the middle portion of the Pliocene Epoch—about three million years ago—is the most recent period when global temperatures were sustained at levels we may see at the end of this century. Three million years may sound like the very distant past but, geologically speaking, it is not. At that time, Earth's continents were in basically the same positions they are now, so ocean circulation patterns (the main vehicle for the transportation of heat) were much the same. Many of the plants and animals that populate our world had already evolved, which makes direct comparisons of living species to their fossil counterparts easy. This symmetry makes the Pliocene a good model of what our future could look like.

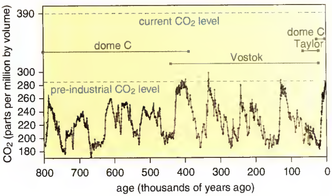
Marci M. Robinson is a micropaleontologist with the U.S. Geological Survey. She worked first with the Pliocene Research Interpretation and Synoptic Mapping (PRISM) project in 1994 as a part-time undergraduate co-op student. Robinson earned her Ph.D. in 2007 at George Mason University, where she studied the coastal geomorphology of the Outer Banks of North Carolina. As a postdoctoral researcher, Robinson surveyed ancient shorelines along the southern and western coasts of Australia and examined the relationship between multiple proxies of ocean paleotemperature. She serves on the advisory panel for the Pliocene Model Intercomparison Project (PlioMIP) and as an assistant editor for the journal Micropaleontology. *Address: Eastern Geology and Paleoclimate Science Center, 926A National Center, Reston, VA 20192. E-mail: mmrobinson@usgs.gov*

Figure 1. Drilling specialists and other crew members aboard the research vessel *JOIDES Resolution* in 2005 prepare a piston core barrel during an Integrated Ocean Drilling Program (IODP) expedition in the North Atlantic. Sediments retrieved on such expeditions are vital to climate studies of the middle part of the Pliocene Epoch, a period 3 million years ago when Earth's temperatures matched those that may occur later this century. A U.S. Geological Survey (USGS) reconstruction of the mid-Pliocene is proving useful to efforts to forecast Earth's future climate. (Photograph courtesy of W. S. Crawford and IODP.)

Atmospheric carbon-dioxide concentrations during the Pliocene were only slightly higher than they are today. Still, global temperatures during the mid-Pliocene were on average 2 to 3 degrees warmer than today, and sea level was on average 25 meters higher. This apparent paradox—a warmer planet with comparable carbon dioxide levels—concerns some researchers who wonder whether even small increases in carbon dioxide can significantly alter our climate.

Soviet scientist Mikhail Budyko pioneered climate studies of the mid-Pliocene with climate change in mind, but he was fueled by optimism rather than alarm. In the 1970s, he concluded that a world warmed by greenhouse gases could benefit regions of the Soviet Union, and he was the first to suggest that using reconstructions of a previous warm period on Earth could be a useful predictor of 21st-century conditions. After encountering Budyko's preliminary data-based climate reconstruction during a U.S. and Soviet scientific exchange, Dick Poore of the U.S. Geological Survey (USGS) and David Rind of NASA concluded that the





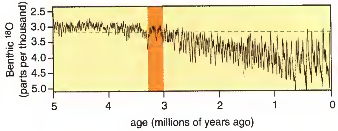


Figure 2. Scientists reconstruct past climate conditions using multiple proxies. At top left, the solid line represents past temperature derived mostly from tree-ring measurements and, after the mid-1800s, from actual measurements. The shaded area covers the range of 21st-century warming forecasts published by the Intergovernmental Panel on Climate Change. Above are CO_2 levels detected in ice cores from three Antarctic locations; the dashed lines indicate pre-industrial and present day CO_2 levels. At left are oxygen isotopic ratios dating back 5 million years. The dashed line represents the modern ratio.

USGS could develop a more quantitative data set for the Pliocene. Since its founding in 1989, the USGS Pliocene Research, Interpretation and Synoptic Mapping (PRISM) project has pursued an unparalleled estimation of what a warmer world looked like.

Reconstructing Climate

A fundamental tenet in geology is uniformitarianism, meaning that the same natural laws and physical processes that operate today were active in the past. By making a few assumptions concerning the stability of ocean chemistry and individual species' ecological tolerances, fossil studies can reconstruct environments of the past at specific locations. When many of these individual reconstructed environments of the same age but from different locations are strung together they give us a picture of global conditions in the past.

This research starts with collecting sediment samples from all over the world. In the deep ocean, most samples come from deep-sea cores drilled by the internationally funded Integrated Ocean Drilling Program (IODP) and its precursors. A European consortium, China, Korea, Australia, New Zealand and India help fund the drilling program, with the U.S. National Science Foundation and Japan's Ministry of Education, Culture, Sports, Science and Technology at the helm. Any researcher can request IODP samples.

IODP expeditions lower coring equipment from research ships to the sea floor and drill down hundreds to thousands of meters to remove ocean sediment containing millions of years of fossil deposits. Several dating techniques allow us to determine how deep in the core to look for Pliocene fossils. (In lakes, the coring process is similar but on a smaller scale. On land, the work can involve deploying hammers and shovels on the side of a road that cuts through a Pliocene deposit.)

Marine cores are usually full of planktic foraminifera, or forams, single-celled organisms that float at or near the ocean surface. A uniquely shaped and ornamented calcium-carbonate shell characterizes each of the nearly 40 species, making it easy to identify them. All are about the size of a grain of sand. Because each species lives in a well-defined range of environmental conditions, the types of species found tells us something about the temperature (or salinity or productivity) at the place and time the forams were alive.

If a fossil assemblage is 60 percent one species and 40 percent another, and if that combination is found in contemporary times where the water temperature is 14 degrees, then you assign a temperature of 14 degrees to that sample. More likely, however, the fossil sample will have as many as 40 species, and that combination will exist in today's ocean over a range of temperatures. The best match to the assemblage may represent a combination of temperatures. A factor analysis is also useful: Modern assemblage data and the associated physical environments undergo a multivariate regression that is applied to the fossil data. That yields an estimate of the temperature in which the assemblage formed.

Most reconstructed ocean temperatures in this research are derived from forams, but other microfossil groups (such as diatoms, radiolaria, ostracods and pollen) are used in much the same way. Another way to estimate paleotemperature is through foram and ostracod shell chemistry. Although made primarily of calcium, carbon and oxygen, they contain a small amount of magnesium. Some magnesium ions replace calcium ions in the calcium carbonate lattice structure. The rate at which this happens is closely tied to the water temperature at the time the shell is secreted. Applied to fossils at the ocean floor as well as those at the surface, this correlation helps us reconstruct deep-ocean thermal gradients. Also useful are alkenones, long-chained organic compounds synthesized by a small group of algae that dwell near the surface of the ocean. The number of double carbon bonds in the chains, or the degree of alkenone unsaturation, varies linearly with the water temperature during growth and provides ocean-temperature estimates that are independent of marine microfossils.

Evidence regarding vegetation, land ice, sea level, deep-ocean temperature, and topography are also in play in the PRISM mid-Pliocene reconstruction.

On land, the distribution of vegetation and land ice is reconstructed through fossil pollen. The extent of land ice on Greenland and Antarctica is loosely defined by the vegetation distribution, but the volume of land ice is closely tied to Pliocene sea level as well. More physical evidence comes from the geological remains of mid-Pliocene shorelines. For instance, the Orangeburg Scarp in the United States, which runs from Florida to Virginia, marks the edge of the Pliocene-era Atlantic Ocean. Today Interstate 95 runs east of the scarp.

Testing the Models

PRISM so far has reconstructed the mid-Pliocene climate with data from 86 marine and 202 land sites. The Pliocene climate coming into view is very different from modern climate in some areas and very similar in others. Temperatures near the equator generally were much like they are today, but temperatures at the poles were much warmer. Moving away from the equator, the temperature difference becomes more extreme.

In the North Atlantic and Arctic oceans, for instance, temperatures at the surface were substantially warmer. These warmer conditions were reflected in the vegetation of Iceland and Greenland, which hosted boreal forest rather than polar tundra vegetation, indicating a Pliocene mean annual temperature at least 10 degrees warmer than what is seen today. In the Southern Hemisphere, Antarctica was vegetated with shrubs along the West Antarctic peninsula and along the coast of Wilkes Land where today there is only ice.

Temperatures in the tropics, for the most part, mimic today's temperatures. In the western tropical Pacific Ocean, near Papua New Guinea, surface water temperatures were similar to present readings—about 29 degrees. In the modern Pacific Ocean, the surface water in the west is a few degrees warmer than in the east, except during an El Niño event when the temperature gradient flattens, an atypical pattern of warmer than normal sea-surface temperatures that alters global weather patterns. But the normal state during the warmest parts of the Pliocene was more like a modern El Niño event. The eastern equatorial Pacific Ocean, near Panama, Colombia and Ecuador, was as warm during the Pliocene as it was in the western equatorial Pacific Ocean, about 1.5 degrees to 2.5 degrees warmer than it is today.

At western continental margins along the coasts of Peru and California, today we find cold, nutrient-rich upwelling zones (and the world's most productive fisheries). Pliocene upwelling was nutrient-rich, judging by paleoproductivity proxies, but as much as 7 degrees warmer. PRISM is incorporating 27 new data sites into its reconstruction, many of them along continental margins, to better understand the nature of the warm upwelling zones of the Pliocene. The Indian Ocean is woefully underrepresented in the current PRISM reconstruction but will come into much sharper view when six new sites are added, as planned. A new focus of research is the historical context of the modern Indian Ocean temperature dipole, another episodic pattern with a strong surface temperature gradient that affects local weather patterns and regional climate.

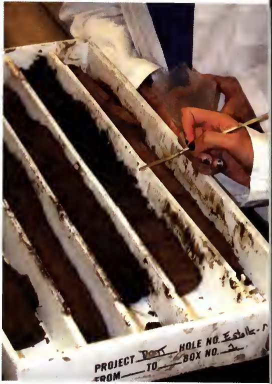


Figure 3. Technicians with the USGS Pliocene Research, Interpretation and Synoptic Mapping (PRISM) Project prepare to remove microfossils from a sediment core. Determining species of the microfossils helped the program bracket the age of depositional features along the Virginia coastal plain. (Photograph courtesy of PRISM.)

Increasingly, climate modelers are testing their accuracy by applying their computational tools to making predictions about the Pliocene climate. And geologists and the modelers are benefitting. For example, in 2004 a significant discrepancy was observed between PRISM's estimates of Pliocene temperatures in the eastern equatorial Pacific and temperatures calculated by the United Kingdom's Hadley Centre Coupled Model (HadCM3). The accuracy of the Hadley model is vital since it has been used extensively for climate prediction and sensitivity studies. In

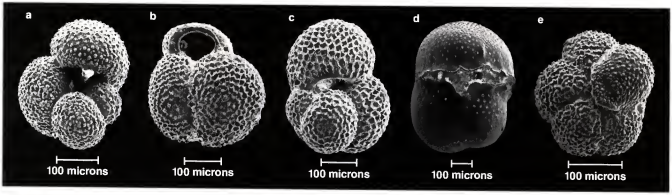


Figure 4. Planktic foraminifera, or forams, are single-celled marine organisms whose shells are preserved in mid-Pliocene sediments. Each species dwelled, as it does today, in a specific environment and has a unique appearance, which makes it easy to identify. The specimens above are a) *Globigerina bulloides*, b) *Globigerinoides ruber*, c) *Globigerinoides sacculifer*, d) *Sphaeroidinella dehiscens* and e) *Turborotalita quinqueloba*. (Scanning electron microscope images courtesy of Marci Robinson.)

its Pliocene simulation of the eastern equatorial Pacific, HadCM3 had concluded that waters west of Panama, Colombia and Ecuador were warmer than they are today. The PRISM reconstruction had not captured that result.

This prompted the author and Harry Dowsett, a micropaleontologist who leads the PRISM group, to search for clues. In IODP's records, we found three cores containing Pliocene sediments in the equatorial region (0 degrees to 6 degrees north or south latitude), one very near the west coast of South America at 84 degrees west longitude, one at 95 degrees west longitude, and one at 110 degrees west longitude. We analyzed the planktic foram assemblages from all three, along with alkenones from the easternmost site that is today within an upwelling region.

The analysis revealed that the eastern equatorial Pacific was in fact warmer during the Pliocene than we'd estimated. The westernmost site showed a 2.5 degrees temperature increase over modern averages. Proxies at the middle site estimated a 2.8 degree temperature increase over modern averages—an incredible agreement between our foram analysis and other data. And at the easternmost site, temperature estimates for the Pliocene were 1.5 and 1.9 degrees warmer than modern measurements. The model projection was correct.

In another example, the PRISM reconstruction observed extreme warmth near Iceland and Svalbard, but the 2004 HadCM3 simulation did not. To address this discrepancy, we started looking into ocean circulation patterns and how the features on the ocean floor in this region have changed since the Pliocene.

North Atlantic deep water forms as cold, salty and more dense surface water sinks just north of Iceland. The Greenland-Scotland Ridge, a bathymetric feature related to the Icelandic hot spot, runs from Greenland to Scotland, through Iceland. This ridge traps the newly formed deepwater as it tries to flow southward along the ocean floor. Geophysical research pioneered by Peter Vogt in the 1970s shows that the Greenland-Scotland Ridge was lower during the Pliocene. Geothermal activity under Iceland has caused the ridge to rise by about 300 meters in the past 3 million years. Additional research by Jim Wright and Ken Miller at Rutgers University showed that changes in the ridge affect deepwater circulation. A pilot study published this year by the author and Dowsett, climate modelers Paul Valdes of the University of Bristol, Alan Haywood of the University of Leeds and Dan Hill of the British Geological Survey, and Steve Jones, a geophysicist at the University of Birmingham, incorporated this change. The adjustment to the model's boundary conditions brought the data and model simulation into agreement.

Comparing data to the model output is essential to improve confidence in climate-model simulations for the future, but these comparisons are in their infancy. Subtle aspects of the comparisons require further attention. We are still addressing some basic differences between the models and the reconstruction, such as how to define "modern" and how to apply seasonal temperature cycles. The data-point–based nature of the reconstruction is also an issue. Both marine and terrestrial cores are scattered unevenly across the globe, clustering in the North Atlantic Ocean and in Western Europe, respectively. Temperature estimates must be extrapolated between points. The biggest data-free space is the middle part of the South Pacific.

In addition, some assumptions are inherent in all proxy data as well as in climate modeling. We assume, for instance, that the ecological tolerances of individual species of foraminifera and the species that produce alkenones haven't changed significantly over time. We assume they lived in the same sorts of environments in the Pliocene as they do now. Likewise, the modelers assume that parameters of climate phenomena, such as cloud formation, can be defined mathematically. Without these assumptions, we wouldn't be able to pursue the science.

Quantifying Uncertainties

Uncertainty in climate projections comes from two sources: the accuracy of the prescribed boundary conditions and the ability of the model to simulate the highly complex global climate system. In this case, the boundary conditions are the individual data sets of the PRISM reconstruction. These are used to set up the model experiments. In the PRISM sea-surface–temperature data set, we minimize uncertainty by utilizing multiple proxies: microfossil assemblage data, magnesium-calcium ratios and alkenones. This approach reduces the overall error compared with results based on a single proxy. It also expands temperature estimation into regions where other proxies are not suitable. For example, planktic foraminifera techniques are not well suited for estimating temperature at high latitudes where foram specimens are often scarce and where assemblages become dominated by a single, often extinct, species. Similarly, alkenone-based techniques are ineffective in the warm tropics because they do not record temperatures above roughly 28 degrees.

Figure 5. The yellow line marks a reconstructed mid-Pliocene shoreline in the southeastern United States. Geochemical and geological data indicate that global sea level was about 25 meters higher three million years ago. This paleoshoreline corresponds with the Orangeburg Scarp, a topographical feature cut by mid-Pliocene wave erosion that stretches from Florida to Virginia.

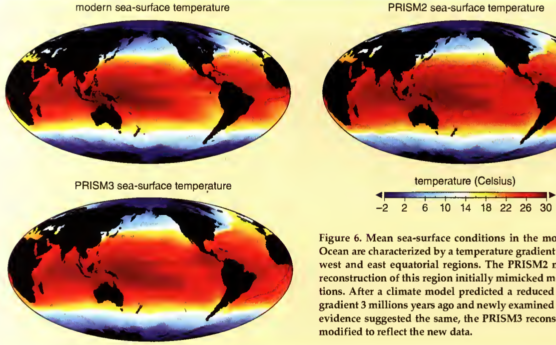


Figure 6. Mean sea-surface conditions in the modern Pacific Ocean are characterized by a temperature gradient between the west and east equatorial regions. The PRISM2 mid-Pliocene reconstruction of this region initially mimicked modern conditions. After a climate model predicted a reduced temperature gradient 3 millions years ago and newly examined geochemical evidence suggested the same, the PRISM3 reconstruction was modified to reflect the new data.

Uncertainties in model simulations can be addressed through model ensembles. In multi-model ensembles, also known as model intercomparisons, different climate models run identical experiments with the same set of boundary conditions and prescribed emissions scenarios, as in the IPCC projections at the beginning of this article.

As the December 2011 deadline approaches for the first draft of the Working Group I Contribution to the IPCC Fifth Assessment Report, the paleoclimate community is gearing up. Model intercomparisons are being completed and analyzed. As part of the Pliocene Model Intercomparison Project (PlioMIP) launched by Dowsett of PRISM, Mark Chandler of Columbia University and the Goddard Institute of Space Studies, and Haywood of the University of Leeds, 17 international climate-modeling groups are currently using the PRISM reconstruction to set boundary conditions in identical experiments. The range in model results will identify model-dependent variability that can then be eliminated from the uncertainty measure. The remaining range will act as error bars to quantify the uncertainty associated with past and future climate simulations.

Another way to bracket model simulation uncertainty is through perturbed physics ensembles. Perturbed physics ensembles use only one model

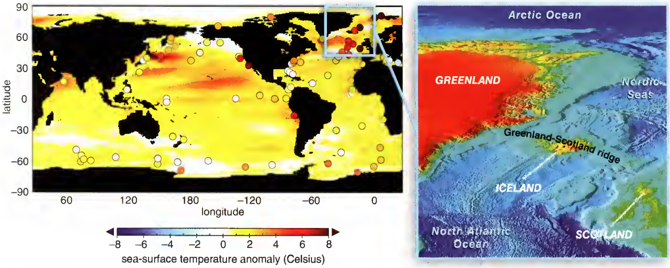


Figure 7. The dots in the map above show mid-Pliocene sea-surface temperature anomalies produced with PRISM data. Estimates for the same period produced by a Hadley Centre Coupled Model simulation are displayed in the color variation on the background. The PRISM temperature estimates near the Greenland-Scotland Ridge in the North Atlantic were warmer than the model-simulated temperatures. The discrepancy prompted a closer look at how geology influences deep water production in that region, which produced a better understanding of how the height of the Greenland-Scotland Ridge likely influences water temperature there.

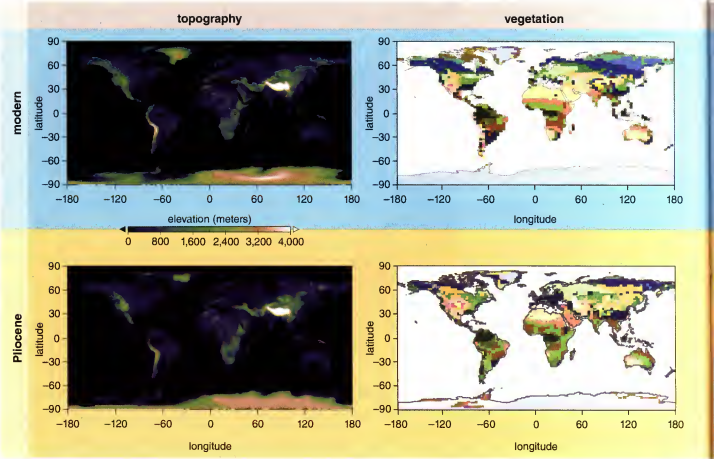


Figure 8. The most recent PRISM mid-Pliocene reconstructions portray significant differences in multiple categories between Earth today and 3 million years ago. In North America, for instance, the northernmost reaches of the Rocky Mountains were lower and the Great Lakes did not exist during the Pliocene. In central and eastern Europe, warm-temperate forests (with subtropical taxa) were abundant. Tropical savannas and woodlands expanded in Africa and Australia. Most of what is tundra on the globe today consisted then of evergreen taiga, a biome characterized by coniferous forests. High-latitude regions were warmer than they are today, and this warming was amplified through sea-ice

at a time and run a series of experiments in which only one variable is changed. The idea is to compare a range of model results, all using different values of that one variable, to the PRISM reconstruction. The best match gives us the most appropriate value for that variable. Thus, that bit of uncertainty is removed.

James Pope, a doctoral student at the University of Leeds, has taken on the task of using a perturbed physics ensemble to estimate uncertainties in Pliocene climate model projections using the HadCM3 and PRISM boundary conditions. This is the first time a perturbed physics ensemble has been applied to a warmer world with high levels of carbon dioxide. In the first experiments, now underway, the variable is climate sensitivity. Climate sensitivity is defined as the global mean annual temperature response to a doubling of atmospheric carbon dioxide. Uncertainty in the value of climate sensitivity is a big part of what creates uncertainty in climate model simulations.

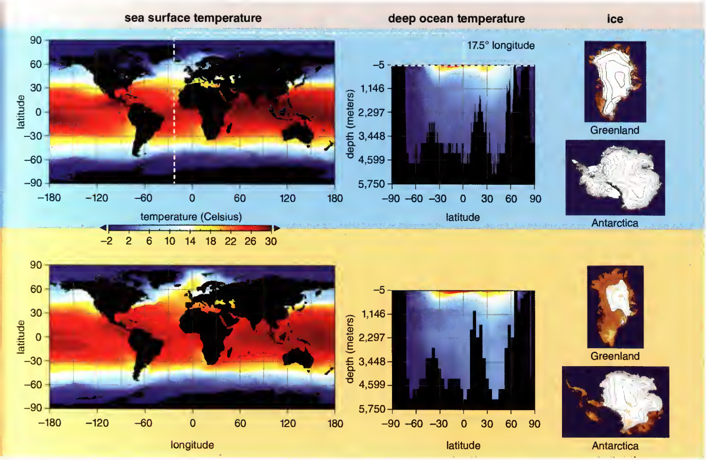
A pilot study by Pope identified low-sensitivity (climate sensitivity = 2.1 degrees) and high-sensitivity (climate sensitivity = 7.1 degrees) model simulations of Pliocene climate as brackets. The high-sensitivity simulation was a close match to the PRISM sea-surface temperatures, but the control, an unperturbed simulation, better captured Pliocene vegetation distribution. These results show promise that a best fit to the data will likely be produced as part of a full ensemble of model simulations.

Making the Best of It

The Pliocene climate is the closest natural analog to the climate we expect near the end of this century and it is an appropriate target for ground-truthing model simulations. But it's important to remember that the Pliocene is unlike the 21st century in key ways. For one, Earth then had not plunged into the glacial-interglacial cycles that have characterized the past few million years. So, rather than trending toward warmer conditions as we are today, the Pliocene was instead the last period before prolonged cooling that led to the ice ages.

More importantly, the Pliocene warm period was characterized by a stable climate state very different from today's harried state of climate disequilibrium. Atmospheric carbon dioxide concentrations, for instance, were relatively stable during the Pliocene, and the climate had adjusted to them. The release of carbon dioxide was natural and gradual, from processes such as volcanic emissions and the decay of plant and animal matter. This is very different from today, as carbon dioxide is released much more rapidly through fossil-fuel use and deforestation. The current climate system is still adjusting and will take a very long time to equilibrate.

In addition, we identified a natural, geologic contributor to Pliocene



feedbacks. Pulses of warm water—about 18 degrees Celsius higher than mid-20th century temperatures there—reached well into the Arctic Ocean, resulting in ice-free conditions for part of the year. North Atlantic deep water, which forms as cold, dense water that sinks north of Iceland and moves southward, was warmer during the Pliocene. The Greenland Ice Sheet and the West Antarctic Ice Sheet were greatly reduced; the East Antarctic Ice Sheet was also smaller. Ice sheet melt then has been associated with a Pliocene global sea level that is equivalent to 22 meters of sea rise today.

warmth that acts on a very slow time scale. The changes in the height of the Greenland-Scotland Ridge, resulting from the movement of magma below Iceland, likely affected the transport of warm surface water into the Arctic Ocean. The warmth seen at these high latitudes was a result of changes that took place on hundred-thousand to million-year timescales. Our climate will not be responding to this particular climate forcing in the near future.

Though the Pliocene is not the perfect analog for the near future, it helps us focus on what may occur. A rise in sea level, for example, and a poleward expansion of tropical temperatures and plants are aspects of the Pliocene that we're likely to see as our climate warms. Earth's high latitudes will likely warm more than the low latitudes, and our ice caps will melt. By how much, we don't yet know. Minor changes in oceanic circulation may have major effects on regional phenomena such as seasonal changes in thermocline depth, where the rate of temperature decrease is the largest, and upwelling. Yes, uniformitarianism is often defined to mean that the present is the key to the past. But we might be wise to remember another accurate description: The past is a key to understanding the future.

Bibliography

Dowsett, H. J., et al. 2011. Sea surface temperatures of the Mid-Piacenzian Warm Period: A comparison of PRISM3 and HadCM3. *Palaeogeography, Palaeoclimatology, Palaeoecology*, in press.

Dowsett, H. J., et al. 2010. The PRISM3D paleoenvironmental reconstruction. *Stratigraphy* 7:123.

Dowsett, H. J., and Robinson, M. M. 2009. Mid-Pliocene equatorial Pacific sea surface temperature reconstruction: A multi-proxy perspective. *Philosophical Transactions of the Royal Society A* 367:109.

Haywood, A. M., et al. 2010. Pliocene Model Intercomparison Project (PlioMIP): Experimental design and boundary conditions (Experiment 1). *Geoscientific Model Development* 3:116.

Haywood, A. M., and Valdes, P. J. 2004. Modelling Pliocene warmth: Contributions of atmosphere, oceans and cryosphere. *Earth and Planetary Science Letters* 218:363.

Intergovernmental Panel on Climate Change. 2007. The physical science basis. *Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Eds. S. Solomon et al. Cambridge: Cambridge University Press.

Robinson, M. M., et al. 2011. Bathymetric controls on Pliocene Arctic sea-surface temperature and deepwater production. *Palaeogeography, Palaeoclimatology, Palaeoecology*, in press.

For relevant Web links, consult this issue of *American Scientist Online:*

http://www.americanscientist.org/issues/id.90/past.aspx

Porphyrins: One Ring in the Colors of Life

A class of pigment molecules binds King George III, vampires and herbicides

Franck E. Dayan and Emilie A. Dayan

One Ring to rule them all,
One Ring to find them,
One Ring to bring them all
and in the darkness bind them.

—J. R. R. Tolkien, *The Fellowship of the Ring*

The porphyrin pathway is ubiquitous in the biological realm, serving throughout the plant and animal kingdoms as the assembly line for the most abundant pigments in nature. The ring-shaped porphyrin molecules bind an array of metal ions, with each combination associated with different biological functions. Chlorophylls bind magnesium to play a pivotal role in photosynthesis. Heme binds iron to coordinate molecular oxygen and carbon-dioxide transport, supports the electron-transport chains necessary for cellular respiration and contributes to the catalytic activities of many enzymes. Porphyrins bind nickel to form coenzyme F_{430}, which plays critically important roles in bacteria that metabolize methane. Vitamin B_{12} is formed from the binding of cobalt to a derivative of porphyrin; lack of the vitamin can result in pernicious anemia and impair the function of the brain and nervous system. Taken together, these porphyrin-derived pigments can be called the "colors of life," in the sense that these rings are necessary to sustain key activities in nearly all organisms.

It should come as no surprise then, that when the porphyrin pathway is disrupted, there can be remarkably far-reaching consequences. Derangements of porphyrin metabolism may have left their marks on the legacy of King George III and the founding of the United States, and on the legend of Dracula. Disruption of porphyrin metabolism can also be useful—for example, in the development of herbicides.

The Key Enzyme

The early porphyrin precursors differ between plants and animals, but they converge at the first committed molecule in the chain of synthesis, delta-aminolevulinic acid. And the subsequent steps involved in the synthesis pathway are universal *(see Figure 2)*. Protoporphyrinogen oxidase is the last enzyme before a major branching point in the synthesis of the "rings of life." The next step finds enzymes catalyzing the binding of various metal ions to protoporphyrin, committing the master ring to the synthesis of an array of biologically important molecules.

In animal cells, protoporphyrinogen oxidase is bound to the inner membrane of mitochondria, the cell's powerhouses. Plants possess two evolutionarily distinct genes, encoded in the nucleus, that create two forms of the enzyme, one found on the outer membrane of chloroplasts (the light-gathering organelles of plant cells) and the other again located in mitochondria. These enzymes have only 25 percent of their amino acids in common and are flanked by different transit peptides, which are short stretches of amino acids that direct newly synthesized proteins to specific locations in the cell.

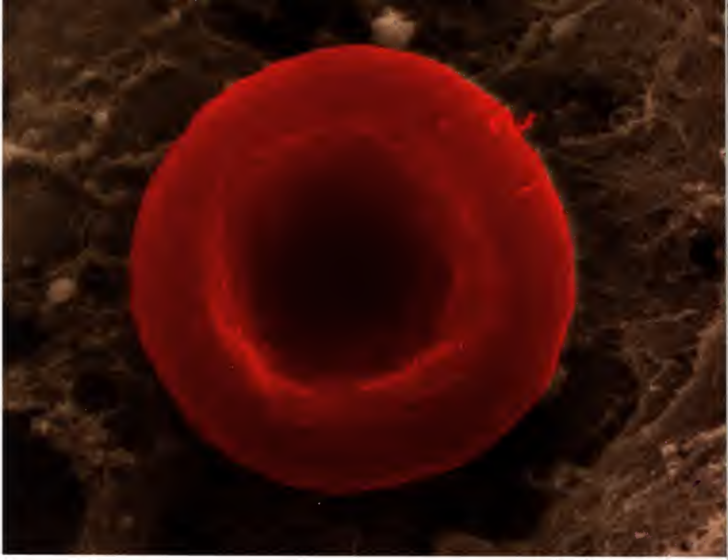
Molecules of a plant's mitochondrial protoporphyrinogen oxidase are loosely associated *(see Figure 3)*. Each molecule, or monomer, possesses three domains: One binds the substrate, the second the membrane and the third an oxidation-reduction factor important in metabolism called flavin adenine dinucleotide (FAD). FAD can accept electrons from upstream metabolites and pass them downstream, making it a key helper in the production of energy in the cell. FAD acts like a roof on the active site. Within the active-site domain, a glycine probably participates in positioning the target substrate, protoporphyrinogen. Some have proposed that another amino acid, arginine, participates in ionic interactions with one of the side chains of protoporphyrinogen, and it may coordinate the orientation of the substrate with respect to the FAD cofactor.

Protoporphyrinogen oxidase catalyzes the removal (or oxidation) of six hydrogen atoms from the colorless substrate to form the photodynamic red pigment protoporphyrin IX. The reaction is not well understood, but it involves three sequential FAD-driven transfers of hydrogen at the central ring of protoporphyrinogen, followed by a complex hydrogen rearrangement. The position of the glycine relative to the opposing FAD ring is critical to the rate of catalysis, as it controls the distance between the carbon ring of the substrate and the FAD cofactor.

The details of the workings of this enzyme pathway are vital to understanding disorders in the production of different porphyrins. Problems with production of heme are particularly significant to humans.

In the human body, heme is synthesized primarily in the liver and bone

Franck E. Dayan is a plant physiologist at the Natural Products Utilization Research Unit of the United States Department of Agriculture and adjunct professor in the Department of Medicinal Chemistry at the University of Mississippi–Oxford. One of his areas of research is the mechanism of action of phytotoxins, and he has published extensively on molecules that interfere with porphyrin synthesis in plants. He received his Ph.D. in Botany from Auburn University in 1995 and has authored 139 scholarly works. Emilie A. Dayan graduated from the Croft Institute for International Studies and the Sally McDonnell Barksdale Honors College at the University of Mississippi–Oxford, where she received a Bachelor of Arts in economics and a Bachelor of Arts in international studies and French. Address for Franck Dayan: USDA-ARS, Natural Products Utilization Research Unit, P.O. Box 8048, University, MS 38677 E-mail: franck.dayan@ars.usda.gov



David M. Phillips/Photo Researchers

Figure 1. The circular shape of a red blood cell (magnified 15,000 times) is reminiscent of the protoporphyrin ring, the pigment molecule that is responsible for its distinctive red color. In addition to their important function in hemoglobin, porphyrins also play key roles in organisms across the biological spectrum.

marrow, and most of it is destined to become bound in the hemoglobin of red blood cells. Accumulation of the purple-red porphyrin intermediates outside of their normal locations is symptomatic of a group of diseases collectively referred to as porphyria. In Greek, *porphuros* means purple. There are genetically inherited dysfunctions associated with each of the steps of heme synthesis shown in Figure 2. Cutaneous porphyrias cause light-dependent swelling and itching of the skin that may develop into rashes and blisters. Acute porphyrias affect the nervous system, in part via deficiency in porphyrin-derived vitamin B_{12}, inducing mental disorders ranging from subtle dysphoria to severe disturbances, as well as engendering pain, muscle numbness and vomiting. Acute porphyria is a term that includes three similar genetically inherited diseases: acute intermittent porphyria, hereditary coproporphyria and variegate porphyria. The development of acute symptoms can be triggered by exposures to foreign substances such as heavy metals, polyhalogenated aromatic hydrocarbons (a type of environmental pollutant) and other compounds that interfere with the porphyrin pathway.

Variegate porphyria is linked to a dysfunction of protoporphyrinogen oxidase. People with this disorder inherit one defective allele, resulting in approximately half the normal activity of protoporphyrinogen oxidase. Patients suffering from variegate porphyria may exhibit both neuropsychiatric symptoms and skin lesions associated with chronic photodermatitis, in addition to increased levels of porphyrins in the feces. Approximately three-quarters of those with the defective allele remain asymptomatic and lead normal healthy lives, but these individuals remain at risk for episodes of illness and must take simple measures to avoid certain triggering factors. Even the few who do become ill usually make a complete recovery and have fewer attacks as they age. However, acute attacks can be very severe in some individuals, and the disorder may affect lives far beyond those who inherit the disease.

The Madness of King George

> The water is of a deeper colour—and leaves a pale blue ring upon the glass near the upper surface.
>
> —Remarks of Sir Henry Halford, one of the attending physicians of King George III, on the monarch's urine (January 6, 1811).

King George III, monarch of England from 1760 until his death in 1820, was one of the longest-serving British rulers. During his reign, the British Empire was well on its way to exerting economic dominance over interna-

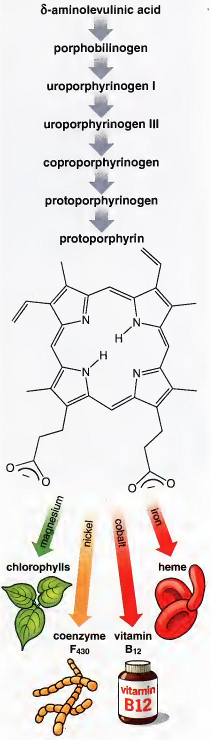


Figure 2. The porphyrin pathway starts with delta-aminolevulinic acid. A different enzyme catalyzes each step in the pathway, with protoporphyrinogen oxidase working at the last step to protoporphyrin. This precursor molecule, whose ring shape is shown in the chemical diagram, can accept different metals to produce pigment molecules essential to metabolism in various organisms. Heme and vitamin B_{12} are vital in animals, chlorophyll in plants, and coenzyme F_{430} in methane-processing bacteria.

tional trade and military supremacy over the oceans. This rise in prominence was, however, tarnished by a strange sickness afflicting the English monarch, whose sometimes irrational decisions may have led to the American Revolutionary War and the subsequent loss of the American colonies.

Royal medical records are replete with well-documented episodes of an incapacitating sickness that plagued King George. In particular, five major episodes each included a prolonged illness in which physical abnormalities were accompanied by profound degradation of his mental faculties. The King's irrational behavior was so debilitating during one episode, which lasted from October 1788 to February 1789, that it instigated a constitutional review known as the Regency Crisis.

These periods of mental incapacity were originally thought to result from a psychiatric illness, but a more precise diagnosis—that the King suffered from porphyria—has been proposed by a number of researchers since the mid-1960s. Indeed, medical records provide evidence that King George exhibited all of the symptoms typically associated with the disease, such as strong abdominal pain, port-wine–colored urine and paralysis in the arms and legs. Initially, it was suggested that the King suffered from acute intermittent porphyria. However, a follow-up study that examined the medical records of forebears, descendants and collateral relatives of King George III led to the diagnosis of variegate porphyria.

King George III may have suffered from a particularly severe form of porphyria. He experienced his first attack in 1765, four years after his marriage to Charlotte of Mecklenburg-Strelitz. Further signs of the disease were recorded by his physicians between 1788 and 1789. From 1811 until his death in 1820, the King's health continued to decline. The disease eventually impaired his faculties to the extent that he required constant care and was confined to his private apartments at Windsor Castle.

Variegate porphyria is an inherited condition, but acute attacks may be triggered by exposure to toxic metals. If King George III did, indeed, carry the genetically defective protoporphyrinogen oxidase, he would have been hypersensitive to the effects of heavy metals and other environmental contaminants common in England during this period. Elemental analysis of the King's hair detected normal to slightly elevated levels of mercury and lead, but levels of arsenic reflecting systemic toxicity were also measured in the King's hair by Timothy J. Cox of the University of Cambridge and his colleagues in 2005. Therefore, arsenic poisoning may have triggered King George III's recurring severe porphyria episodes.

In 1968 British physician Ida MacAlpine and her colleagues studied the medical records of 13 generations of members of King George III's lineage spanning more than 400 years. They found biochemical evidence that the King and numerous other members of the royal line may have suffered from porphyria *(see Figure 4)*. Several other family members may have been carriers of the genetic defect, but their medical records do not suggest that they exhibited any symptoms of porphyria. This lack of evidence is not surprising because variegate porphyria is a low-penetrance disorder, and 75 percent of carriers are asymptomatic. Subsequent studies in 1982 by Lindsay C. Hurst of Moorhaven Hospital in England have suggested that porphyria can be traced back even further to Henry VI of England and Charles VI of France in the 14th and 15th centuries. A 1996 study by Martin J. Warren, of the University of Kent, and his colleagues looked at more recent generations and found a link to the disease in the descendants of King George III.

Porphyria's grip on King George affected both his physical and mental health, which may have had a tremendous impact on the course of history. Following the French and Indian War, which ended in 1763, Britain levied a series of taxes and enacted several laws to strengthen its authority over its American colony. The colonists, who were British subjects, objected to this assertion of power and famously complained that there should be "no taxation without representation." As political tensions heightened, the King took increasingly strident positions, such as sending British troops to squash the Boston tea party rebellion. He also removed local governments and installed royal officials in their stead. Patriots mobilized their militias, and the American Revolutionary War started in 1775. Historical accounts show that the King obstinately (perhaps even irrationally) resisted any diplomatic solutions that were pursued throughout

the duration of the war. This outcome eventually led the colonists to declare independence from British rule in July 1776. The text of the Declaration of Independence itemizes King George's unwillingness to compromise:

> Such has been the patient sufferance of these Colonies; and such is now the necessity which constrains them to alter their former Systems of Government. The history of the present King of Great Britain is a history of repeated injuries and usurpations, all having in direct object the establishment of an absolute Tyranny over these States.

The war essentially ended with the surrender of Lord Cornwallis at Yorktown on October 19, 1781, followed by formal British abandonment of any claims to the United States with the Treaty of Paris in 1783.

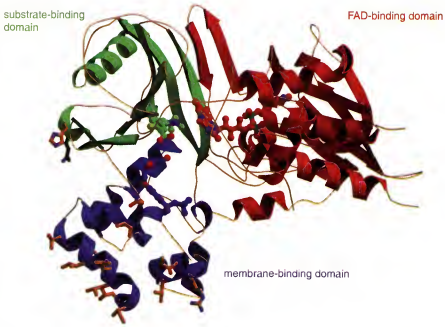


Figure 3. The enzyme protoporphyrinogen oxidase consists of three domains. The membrane-binding domain *(blue)* attaches the enzyme to the cellular organelle where the enzyme carries out its function. The substrate-binding domain *(green)* joins the enzyme to the molecule on which is it is working, in this case, protoporphyrinogen. The FAD-binding domain *(red)* hosts an oxidation-reduction factor called flavin adenine dinucleotide (FAD). The enzyme, dependent on FAD, removes hydrogen atoms from protoporphyrinogen to produce protoporphyrin. (Image courtesy of Michael Koch, 2004, EMBO Journal.)

Certain historians have postulated that porphyria paralyzed the energies of King George III, clouding his judgment and making him temperamental. Such views have also been popularized in the film *The Madness of King George*. Mood swings and impaired reason led to a number of unwise and rash decisions that would ultimately sever the ties to British subjects living in the American colonies, to the detriment of both Britain and loyalists in America. These historians conjecture that had King George III not been incapacitated by illness, he might have been able to avoid the American Revolutionary War.

It should be noted, however, that despite the medical records suggesting that the King suffered from recurrent attacks of porphyria, this retrospective diagnosis is still the subject of dispute. Two studies, published in 2010 by Timothy J. Peters at the University of Birmingham and his colleagues, reexamined the King's medical records from clinical and psychiatric perspectives, and they concluded that he might have instead suffered manic-depressive psychosis. The source of the King's malady may never be conclusively proven, but there are other interesting stories that may have been marked by porphyria.

The Art Archive / Library of Congress / Superstock

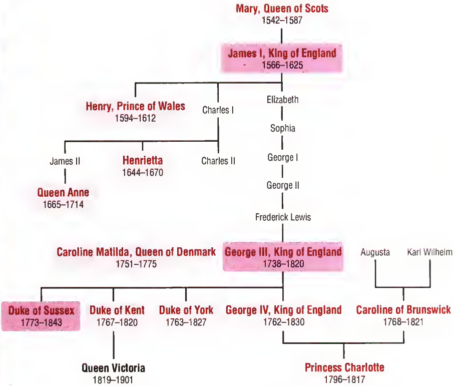
Figure 4. This lithograph, published in 1845 by Nathaniel Currier and James Merritt Ives, shows an idealized depiction of the surrender of Lord Cornwallis of the British military (symbolized by the traditional handing over of his sword) to George Washington, after Cornwallis's defeat at Yorktown, Virginia, in 1781. Porphyria in King George III may have hindered him from considering diplomatic solutions that could have prevented the American Revolutionary War.

Vlad Tepes and Dracula

> Then, with the fear on me of what might be, I drew a ring so big for her comfort, round where Madam Mina sat. And over the ring I passed some of the wafer, and I broke it fine so that all was well guarded. She sat still all the time, so still as one dead.
>
> —Bram Stoker, *Dracula*

Vlad III, born in 1431, was Prince of Wallachia (a region of Romania) during the mid- to late-15th century. A descendent of Mircea the Great, King of Wallachia from 1386 until 1418, Vlad was Knight of the Order of the Dragon, a religious and chivalric order created in Hungary whose purpose was to protect the interests of Catholicism and the Holy Roman Empire, especially against the Ottoman Empire. Vlad was an authoritarian ruler who built a strong military organization that permitted increased and safer trade with neighboring countries.

Vlad was extremely cruel to his foes, whether fellow countrymen or soldiers of invading armies. On Easter Sunday 1459, he arrested and impaled many of his nobles who had rebelled and killed his father and his brother. His prefer-



©Samuel Goldwyn/Everett Collection

Figure 5. The movie *The Madness of King George* dramatized the affliction of the British monarch, portrayed here by actor Nigel Hawthorne *(left)*. Research by Ida MacAlpine and her colleagues in 1968 documented members of English royal family who seem to have suffered from or transmitted porphyria *(right)*. Names in red are individuals whose medical records show clinical evidence of the disease. Darker, purple backgrounds indicate individuals who exhibited discolored urine. Because 75 percent of carriers of variegate porphyria are asymptomatic, it is not surprising that a large number of royal family members may have not shown signs of porphyria, even if they transferred the allele for the disease to their offspring. (Chart adapted from MacAlpine et al, 1968.)

ence for this form of punishment consequently earned him the nickname Tepes, which means "the impaler." Vlad Tepes's appetite for torturing his enemies seems to have been bottomless. He condemned people to death using a myriad of gruesome methods such as skinning, decapitation, hacking, strangulation, hanging, boiling and burning. Other punishments included cutting off noses, ears, sexual organs and limbs. Vlad Tepes is said to have killed 20,000 to 40,000 European civilians, mostly by impalation.

Having been captured and tortured by the Ottoman Empire during his childhood, Vlad was particularly ruthless toward invading Muslim armies. In 1462, when Sultan Mehmed the Conqueror pursued Vlad, he encountered a "Forest of the Impaled" as he entered the capital of Wallachia, where thousands of stakes held the rotting cadavers of 20,000 Turkish captives. In all, Vlad Tepes is said to have impaled as many as 100,000 Turkish Muslims.

Although his bloodthirstiness is not known to have extended to literal ingestion of his enemies' blood, Vlad Tepes also served as the model for an enduring fictional villain, Dracula. Mythological "undead" beings who feed on the blood of the living have been recorded in most cultures from the earliest times. The term *vampire*, however, was not developed in Western Europe until Eastern European vampire legends were popularized during the early 18th century. Bram Stoker, an Irish

Universal/The Kobal Collection

Figure 6. Vlad Tepes used the name Dracula, meaning "son of the dragon" *(right)*. His bloodthirstiness was combined with vampire stories to create a character of the same name (played by Bela Lugosi in 1931, at left). Early vampire myths may have roots in porphyria, which produces similar physical symptoms.

writer, first heard of Vlad Tepes in 1890 during an encounter with the Hungarian professor Ármin Vámbéry. Stoker became fascinated by the prince's perplexing personality and developed his phantasmagoric fictional character of Dracula the vampire from the vivid and gruesome historical accounts of Vlad's cruelty. He worked on his novel for several years and published *Dracula* in May 1897. Stoker's choice of the name Dracula is also extracted from the life of the prince of Wallachia. Vlad and his father were both Knights of the Order of the Dragon. In the Romanian language, "Dragon" is pronounced "Dracul," and the noble families of Romania called Vlad's father Dracul. Dracula simply means "the son of Dracul," and Vlad ultimately adopted this as a nickname.

There is no evidence that Stoker knew about porphyria or the striking similarity between the appearances of the character he created and some of the symptoms commonly associated with the disease. (There is also no indication that Vlad Tepes suffered from porphyria.) However, David Dolphin, a Canadian chemist who edited an authoritative seven-volume series on porphyrins, realized the similarity between the imaginary character of Dracula and some of the symptoms of patients affected by porphyria.

Dolphin postulated that these stories may have been derived from the mythologization of actual individuals affected by porphyria. Indeed, the accumulation of photodynamic pigments in the epidermis and retina renders porphyria victims exceedingly sensitive to light. Such harmful effects can lead to disfigurations such as the withering of fingers and lips, and gums may tighten to reveal fanglike teeth with reddish hues due to elevated porphyrin levels.

Although it is possible that a dysfunction in protoporphyrinogen oxidase may prove to be a scientific basis that contributed to the creation of the myth of vampires, it is important to remember that patients afflicted with porphyria are by no means vampires and that it is certainly an unfortunate coincidence that this serious malady has been connected to such a dreadful, fictitious creature. No one suffering from porphyria deserves a rendezvous with Buffy the vampire slayer. Instead, its victims should be able to seek medical attention and not feel stigmatized.

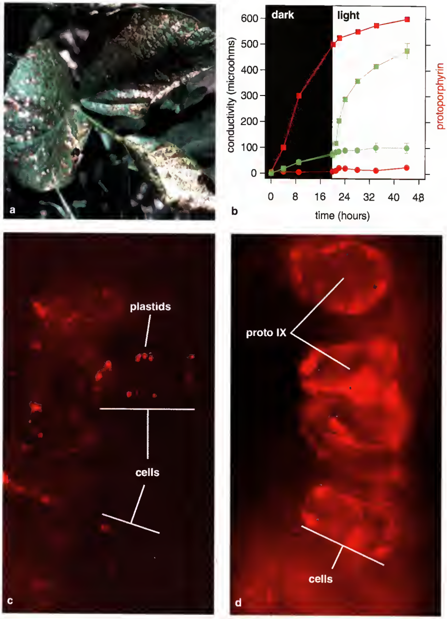


Figure 7. Herbicides that inhibit the enzyme protoporphyrinogen oxidase cause photobleaching in weeds *(a)*. A graph *(b)* shows how protoporphyrin levels *(red)* and electrolyte leakage *(green)* in leaf tissue change from darkness to light exposure in the absence *(circles)* and presence *(squares)* of the herbicide. Cucumber cotyledons (embryonic leaves) normally show protoporphyrin accumulation only in cellular organelles called plastids *(c)*. However, three hours after cotyledons were treated with the herbicide, their cells show a buildup of protoporphyrin (proto IX) throughout the cell *(d)*. (Top left image courtesy of the author; bottom two images courtesy of Stephen O. Duke, USDA.)

Porphyrias in Plants

> Thy design I know well,
> But little I care:
> Who wins the ring
> Will rule by its might.
>
> —Richard Wagner, *The Ring of the Nibelung*

Porphyrias in plants have largely gone unnoticed because dysfunction in the highly regulated porphyrin pathway probably carries a serious fitness penalty, and affected plants most likely do not survive. Gongshe Hu and colleagues at the University of Missouri have, however, recently characterized a maize line called Les22 exhibiting light-dependent lesions on its foliage. Les22 was cloned using a mutator-tagging technique, and analysis revealed that it encodes uroporphyrinogen decarboxylase, a key enzyme in the biosynthetic pathway

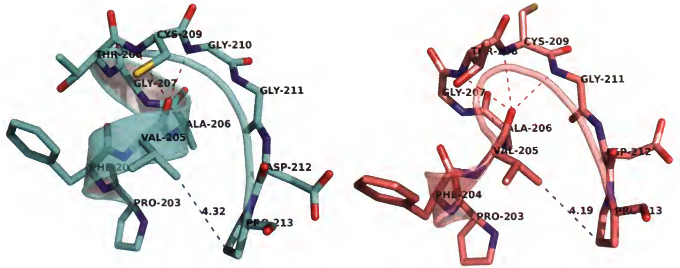


Figure 8. Weed resistance to herbicides that inhibit protoporphyrinogen oxidase has been slow to evolve, but a resistant type of water hemp has recently been identified. The cause of the resistance has been traced to a gene called PPX2L. The normal enzyme is shown at left, the modified enzyme is at right. The modified PPX2L gene has a codon deletion for the amino acid glycine labeled GLY-210, shown near the top right of the blue molecule. This deletion decreases the catalytic efficiency of the enzyme, and the lower performance is compensated for by a change in the binding mode of the herbicides, significantly decreasing their affinity for the enzyme. (Images courtesy of the author.)

of chlorophyll and heme. Mutations in uroporphyrinogen decarboxylase are associated with porphyria cutanea tarda in humans, which manifests itself as skin injury resulting from an excessive accumulation of uroporphyrin. In his editorial about the study, Crispin B. Taylor of the American Society of Plant Biologists referred to this plant as a vampire plant because its phenotype is similar to what is observed in photosensitized patients afflicted with porphyria. To date, plants with other natural dysfunctions in their porphyrin pathway, such as a deficiency in protoporphyrinogen oxidase as in variegate porphyria, are yet to be identified.

The first herbicide that included a protoporphyrinogen-oxidase inhibitor, nitrofen, was introduced in 1964, but the mechanism of action of these herbicides perplexed scientists for many years. They cause rapid photobleaching and light-dependent desiccation of foliage, resulting in leaf cupping, crinkling, bronzing and necrosis. The lesions on the foliage are caused by a loss of membrane integrity that leads to cellular leakage. Other physiological responses include inhibition of photosynthesis; evolution of ethylene, ethane and malondialdehyde; and, finally, bleaching of chloroplast pigments. Protoporphyrinogen-oxidase inhibitors are known to cause temporary injury to the foliage of treated crops they were designed to protect, but these plants recover rapidly because they quickly degrade the herbicide.

The light-dependent activity of protoporphyrinogen-oxidase inhibitors was recognized by 1969, but their actual mechanism of action eluded scientists for decades. In the late 1980s, the French research group of Michel Matringe and his colleagues found that the loss of membrane integrity results from the accumulation of a photodynamic pigment, which was soon identified as protoporphyrin *(see Figure 7)*. Although this result pointed to the inhibition of a biochemical step downstream from the point of synthesis of protoporphyrin, none of the enzymes tested proved to be the target. The same research group, in collaboration with experts in porphyrin biochemistry, realized how similar this effect is to the accumulation of protoporphyrin in human deficiency of protoporphyrinogen-oxidase activity (variegate porphyria) and reported that these herbicides indeed targeted the chloroplastic enzyme. The apparent paradox of inhibition of an enzyme leading to the accumulation of its catalytic product is explained by the altered subcellular compartmentalization of porphyrin intermediates. As with humans with variegate porphyria, inhibition of protoporphyrinogen oxidase induces an uncontrolled accumulation of protoporphyrinogen, which in plants leaks out of the chloroplasts' outer membrane into the cytoplasm, where it is converted into the highly photodynamic protoporphyrin. In the presence of light, protoporphyrin generates highly reactive singlet oxygen species that cause lipid peroxidation of the relatively unprotected plasma membrane.

Resistance to protoporphyrinogen-oxidase–inhibiting herbicides in weeds has been slow to evolve, but resistant biotypes of water hemp *(Amaranthus tuberculatus)* have recently been identified. These biotypes have broad levels of cross resistance to several classes of protoporphyrinogen-oxidase inhibitors. Resistance has been associated with reduced accumulation of protoporphyrin, which was accompanied by reduced membrane damage. Resistance was characterized as an incompletely dominant trait on a nuclear gene called PPX2L. This gene encodes the mitochondrial protoporphyrinogen-oxidase homolog, which is not the primary target of these inhibitors. However, this gene possesses an amino-terminal extension, which leads to dual targeting of the gene product toward both plastids and mitochondria.

Instead of a usual single-point mutation resulting in the substitution of one amino acid for another, the research group of Patrick J. Tranel at the University of Illinois discovered that the herbicide-resistant PPX2L gene has a codon deletion for glycine at a critical position *(see Figure 8)*. This improbable mutation may have occurred as a result of the slippage of DNA poly-

merase in an array of microsatellite repeats. One of us (Franck E. Dayan) and his colleagues established that the deletion did not affect the affinity of protoporphyrinogen or the FAD content, but it decreased the catalytic efficiency of the enzyme. The suboptimal efficiency was compensated for by a change in the binding mode of the herbicides and a significant increase in their affinity for protoporphyrinogen oxidase. The seemingly innocuous deletion of a glycine in fact altered the conformation of protoporphyrinogen oxidase significantly because this amino acid residue plays a key role in stabilizing the helix.

One Ring

I saw eternity the other night
Like a great ring of pure and
  endless light,
All calm as it was bright,
And round beneath it time in
  hours, days, years,
Driven by the spheres,
Like a vast shadow moved in
  which the world
And all her train were hurled.

— Henry Vaughan, *The World*

Nature's propensity to embrace specific molecules (amino acids, nucleic acids, fatty acids and carbohydrates) as the building blocks of structures with grander purposes seems to be consistent with its usage of the porphyrin ring. The unique metabolic pathway producing protoporphyrin, the master ring from which the other "rings of life" are derived, is ubiquitous and appeared early in the evolution of life. The biological importance of these molecules dictates that any disturbance of their synthesis, and particularly of the terminal step that forges the "One Ring to rule them all," has catastrophic consequences in all aspects of life.

References

Boulton, J. 2000. *The Knights of the Crown: The Monarchical Orders of Knighthood in Later Medieval Europe 1325-1520*. Suffolk, UK: Boydell Press.

Cox, T. M., N. Jack, S. Lofthouse, J. Watling, J. Haines and M. J. Warren. 2005. King George III and porphyria: An elemental hypothesis and investigation. *Lancet* 366:332.

Dayan, F. E., P. R. Daga, S. O. Duke, R. M. Lee, P. J. Tranel and R. J. Doerksen. 2010. Biochemical and structural consequences of a glycine deletion in the α-8 helix of protoporphyrinogen oxidase. *Biochimica et Biophysica Acta* 1804:1548.

Dolphin, D. 1978–1979. *The Porphyrins, Volumes 1-7*. New York: Academic Press.

Dolphin, D. 1985. Werewolves and vampires. *Annual meeting of American Association for the Advancement of Science.*

Heinemann, I. U., M. Jahn and D. Jahn. 2008. The biochemistry of heme biosynthesis. *Archives of Biochemistry and Biophysics* 474:238.

Hu, G., N. Yalpani, S. P. Briggs and G. S. Johal. 1998. A porphyrin pathway impairment is responsible for the phenotype of a dominant disease lesion mimic mutant of maize. *Plant Cell* 10:1095.

Hurst, L. C. 1982. Porphyria revisited. *Medical History* 26:179.

Koch, M., C. Breithaupt, R. Kiefersauer, J. Freigang, R. Huber and A. Messerschmidt. 2004. Crystal structure of protoporphyrinogen IX oxidase: a key enzyme in haem and chlorophyll biosynthesis. *EMBO Journal* 23:1720.

Layer, G., J. Reichelt, D. Jahn and D. W. Heinz. 2010. Structure and function of enzymes in heme biosynthesis. *Protein Science* 19:1137.

MacAlpine, I. and R. Hunter. 1966. The "insanity" of King George III: A classic case of porphyria. *British Medical Journal* 1:65.

MacAlpine, I., R. Hunter and C. Rimington. 1968. Porphyria in the royal houses of Stuart, Hanover, and Prussia: A follow-up study of George III's illness. *British Medical Journal* 1:7.

MacAlpine, I. and R. Hunter. 1969. *George III and the Mad Business.* New York: Pantheon Books.

Matringe, M., J.-M. Camadro, P. Labbe and R. Scalla. 1989. Protoporphyrinogen oxidase as a molecular target for diphenyl ether herbicides. *Biochemical Journal* 260:231.

Mochizuki, N., R. Tanaka, B. Grimm, T. Masuda, M. Michael, A. G. Smith, A. Tanaka and M. J. Terry. 2010. The cell biology of tetrapyrroles: a life and death struggle. *Trends in Plant Sciences* 15:488.

Patzoldt, W. L., A. G. Hager, J. S. McCormick and P. Tranel. 2006. A codon deletion confers resistance to herbicides inhibiting protoporphyrinogen oxidase. *Proceedings of the National Academy of Sciences of the U.S.A.* 103:12329.

Peters, T. J., and D. Wilkinson. 2010. King George III and porphyria: A clinical re-examination of the historical evidence. *History of Psychiatry* 21:3.

Peters, T. J., and A. Beveridge. 2010. The madness of King George III: A psychiatric re-assessment. *History of Psychiatry* 21:20.

Poblete-Gutierrez, P., T. Wiederholt, H. F. Merk and J. Frank. 2006. The porphyrias: clinical presentation, diagnosis and treatment. *European Journal of Dermatology* 16:230.

Röhl, J. C. G., M. J. Warren and D. M. Hunt. 1998. *Purple Secret: Genes, "Madness" and the Royal Houses of Europe*. London: Bantam Press.

Schulze, J. O., W.-D. Schubert, J. Moser, D. Jahn and D. W. Heinz. 2006. Evolutionary relationship between initial enzymes of tetrapyrrole biosynthesis. *Journal of Molecular Biology* 358:1212.

Taylor, C. B. 1998. Vampire plants? *Plant Cell* 10:1071.

Thunell, S. 2000. Porphyrins, porphyrin metabolism and porphyrias. I. Update. *Scandinavian Journal of Clinical Laboratory Investigation* 60:509.

Warren, M. J., M. Jay, D. M. Hunt, G. H. Elder and J. C. G. Röhl. 1996. The maddening business of King George III and porphyria. *Trends in Biochemical Science* 21:229.

For relevant Web links, consult this issue of *American Scientist Online:*

http://www.americanscientist.org/issues/id.90/past.aspx

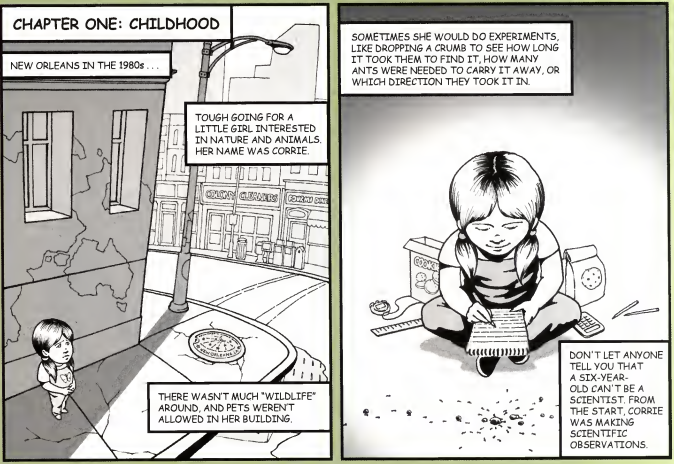
"He thinks he should have been credited in your paper because you used his computer through the distributed computing network."

Sightings

Serious Science, Comic-Book Style

More than 300 live harvester ants, *Pogonomyrmex occidentalis*, are on display in the ant farm that welcomes visitors to the The Field Museum temporary exhibit The Romance of Ants. Photograph by Karen Bean.

The people who create museum exhibits strive to grab attention. That's not so simple when budgets have slimmed, but visitors' expectations have remained super-sized. At The Field Museum in Chicago, exhibition development director Matt Matcuk and his team recently found one way. While assembling the temporary exhibit The Romance of Ants, they stuck to some fundamentals: the universal love of story and people's inherent interest in others. They also made it fresh by mixing media, including a comic-book style narrative and museum-grade photographs by University of Illinois biologist Alex Wild. A passion for science is conveyed through the real-life journey of Corrie Moreau, an entomologist and a museum assistant curator. Alexandra Westrich, an artist and aspiring entomologist working in Moreau's laboratory, created the artwork. The exhibit, including the edited portion shown here, will be on view in Chicago through 2011. Moreau and Westrich described their backgrounds and this nontraditional project to American Scientist *associate editor Catherine Clabby.*



Almost every ant you ever see is female, like these *Pseudomyrmex spinicola* worker ants seen on the tip of an Acacia tree thorn. Photograph by Alex Wild (www.alexanderwild.com).

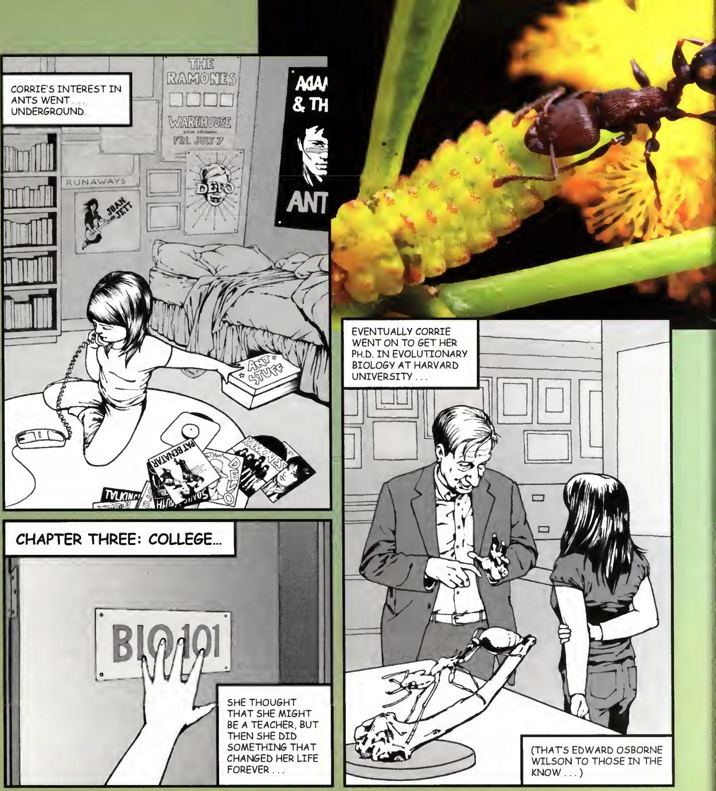
The Scientist

When I was first approached about the exhibit I was very interested in sharing my science. When I learned that the team planned to tell much of the story using a graphic novel format, I was caught a bit off guard. I am quite comfortable sharing my science, but initially felt a bit uncomfortable about the exhibit being about my journey. In the end it became as much about my journey as it was about highlighting the amazing insects I work on and my scientific research.

Growing up in an urban environment meant that there were not a lot of wild creatures around, outside of ants and other insects. I think that is part of the reason I was so drawn to them. You can find ants and insects anywhere. This being said, it is not all that "cool" for a teenage girl to like bugs and science. So I gave in to peer pressure and let those passions go underground (so to speak) for a few years.

During my junior high and high school years I became very interested in environmental issues and the idea of seeing new parts of the world. This led to me to move to California for college, where I chose San Francisco State University because it had a biology major that focused in entomology. It was the perfect fit. One of the reasons I enjoy science is that each discovery can take you in a new direction. For my Ph.D., I wanted to ask new questions about ant evolution. This led me to work with Edward O. Wilson and Naomi Pierce at Harvard University. The knowledge and inspiration I gained from them was amazing.

It wasn't until the exhibit had been open for a few weeks that I realized the power of using a cartoon narrative. I started to get emails from public visitors, children and adults. If the display inspires others to consider a career in science or teaches them that natural history museums are active research institutions, then the exhibit has been a success in my book.



The Artist

I've been a committed enthusiast of both nature and art since childhood. While a student at the School of the Art Institute of Chicago, I favored science electives such as "The Insect World" and "Animal Behavior" over studio classes. This was when I started to seriously consider applying my artistic abilities to the sciences, first through very basic avenues such as scientific illustration, but gradually extending to alternative media (such as comics and "zines") that could appeal to a wider audience.

I first met Corrie in 2009, when I started volunteering at The Field Museum. After a few months, Corrie offered me a position as a collection assistant. My work includes point mounting, documenting and imaging ants that she collects in the Florida Keys as part of a long-term survey of the diversity of ant communities (both native and exotic) unique to these islands.

For this project, Matt Matcuk conducted the initial interviews and wrote the actual narrative. Many of the highlights of Corrie's biography (particularly her formative years in New Orleans) were new to me. I think some of the most powerful events in the narrative happen in the beginning, when Corrie's fascination with the natural

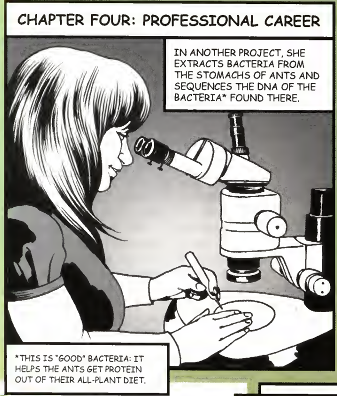
This *Podomyrma adelaidae* ant shadows a lycaenid caterpillar, which secretes a liquid useful to ants. In return, the ant protects the caterpillar from parasites. Photograph by Alex Wild (www.alexanderwild.com).

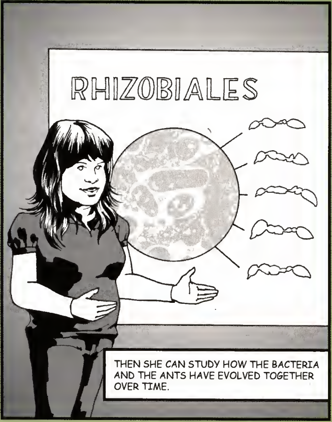
world first takes shape. The evolution of her interest in science—from rapt observations about ant behavior as a little girl to award-winning science projects—provides inspiring insight into the endurance of dreams and the small but purposeful steps that lead to their fulfillment.

I hope The Romance of Ants draws attention to the research and curation efforts of individuals such as Corrie who make up The Field Museum's sizeable (if largely unseen) research and collections staff. I hope it also motivates younger visitors to take an interest in science. Corrie's story is enormously inspirational and, with any luck, will mobilize kids to join the profession. (Honestly, the world can never have too many entomologists).

A fuller look at The Romance of Ants and its graphic narrative is available here: http://romanceofants.com.

In Sightings, American Scientist *publishes examples of innovative scientific imaging from diverse research fields.*





The leafcutter ant *Atta texana* can lift up to 20 times its own weight. Photograph by Alex Wild (www.alexanderwild.com).

Galileo's Discoveries After 400 Years

Noel M. Swerdlow

GALILEO. J. L. Heilbron. xiv + 508 pp. Oxford University Press, 2010. $34.95.

GALILEO: Watcher of the Skies. David Wootton. Yale University Press, 2010. xii + 328 pp. $35.

The year 2010 marked the 400th anniversary of Galileo's discoveries with the telescope, and the event and its author were commemorated with conferences, lectures, exhibits and publications all over the world. The literature on Galileo already exceeds in quantity the literature on every other scientist. (His closest competitors are Einstein and Darwin; even Newton falls far behind.) Why should this be? There must be a dozen scientists whose contribution is of greater importance. Also, much of Galileo's own work is defective, some of it in ways he was himself aware of. And yet this attention to Galileo is not misplaced and has every promise of continuing for as long as the history of science and science itself are pursued. Consider Galileo's life and work.

Galileo Galilei was born in 1564, the son of Vincenzio Galilei, a Florentine cloth merchant and musician best known for his theoretical writings on ancient and modern music. From Vincenzio, Galileo learned to play the lute and sing, to discover by observation and experience, and to be skeptical of authority and the supernatural. His father wished him to study medicine, to which end he attended the University of Pisa, but he greatly preferred mathematics, which he studied outside the university. He left in 1585 without a degree, for four years taught mathematics privately, and in 1589 received an appointment teaching mathematics at Pisa even though he was still without a degree. Three years later he received a call to lecture on mathematics at the *Studio di Padova*, then the most distinguished university in Europe. That Galileo, who had as yet published nothing, was appointed to such a post can only show recognition of his greatest quality, his genius.

His principal interest was mechanics—the application of mathematics to statics, kinematics and dynamics—and he understood that the object of mechanics is to reduce these subjects to geometry, without the need for hidden forces. During the 18 years he spent in Padua, he made his most important discoveries in mechanics, both experimental and theoretical: the acceleration of a falling body, the parabolic trajectory of a projectile and the resistance of solids to fracture. Up to this point in his life, Galileo's concern with astronomy was slight, although he did compute any number of horoscopes. But when Johannes Kepler sent him a copy of his *Mysterium cosmographicum* in 1597, the first important successor to Nicolaus Copernicus's *De revolutionibus* of 1543, Galileo wrote back that he had himself arrived at the Copernican opinion many years ago and from that assumption had discovered the cause of many natural effects. Kepler guessed, correctly, that Galileo was referring to the tides.

Through all his years in Padua Galileo published little and was barely known outside his circle of friends and students in Padua, Venice and Florence. Then everything changed. In the early summer of 1609, he learned of an optical device recently invented in the Netherlands, which made distant objects appear close. He immediately figured out how to make such a thing from spectacle lenses—a "spyglass" giving an upright image of 2 or 3 times magnification, which he soon improved to 8 or 9 times. In September he offered it to the Republic of Venice as a military secret for spotting distant ships at sea, and he was rewarded with a lifetime appointment at Padua and a doubling of his salary.

But he had bigger things in mind. By December he had made an instrument of 20 times magnification and soon after that 30, and he began looking to the heavens. In two months he made more discoveries that changed the world than anyone before or since. He discovered, and confirmed, that the surface of the moon, of which he made strikingly realistic drawings, was rough and mountainous. He found that there are a vast number of stars that are smaller—we would say fainter—than those visible to the un-

aided eye, and saw that the Milky Way is made up of stars beyond counting. Most surprising of all, he saw that Jupiter is accompanied by four small stars that move around it. (Kepler soon introduced the term *satellites*—from *satelles*, an attendant upon an important person—to refer to them.) The discovery of Jupiter's little stars made Galileo resolve to publish quickly, before someone else had the bright idea of turning the new optical device on Jupiter. He recorded nightly observations of the configurations of Jupiter and its companions from January 7 to March 2, and by March 13, 1610, published his new discoveries in the *Sidereal Messenger*. The little book was soon known throughout Europe, and Galileo became the most famous natural philosopher in the world.

In the course of that year Galileo went on to discover the peculiar shape of Saturn, which was correctly explained as a ring nearly 50 years later by Christiaan Huygens; the phases of Venus, showing that it must move about the Sun and shine by reflected sunlight; and spots on the Sun that move and appear and disappear, showing that the Sun rotates and the heavens are subject to change. And that is why after 400 years, we are having conferences, lectures, exhibits and publications, and it is one reason that Galileo is still the most famous natural philosopher in the world.

But there are other reasons. The *Sidereal Messenger* is Copernican—Galileo states that he will prove that the Earth is a planet in an intended work to be titled *System of the World*—and so too are the *Letters on Sunspots* of 1613, in which he reports his later discoveries. By the summer of 1610 he had left Padua to take the positions of Mathematician and Philosopher to the Grand Duke of Tuscany in Florence and Chief Mathematician of the University of Pisa, with no teaching responsibilities. He had all he wanted, but he had made what turned out to be a serious mistake, for when trouble came, the Grand Duke could not provide the protection he would have received in Padua from the Republic of Venice.

In December 1613, in answering a letter from his former student Benedetto Castelli concerning scriptural objections to Copernican theory, Galileo set out his own views of Scripture and science, offering an ingenious interpretation of Joshua's making the Sun stand still to show that not only does Holy Scripture not oppose Copernican theory, it actually supports it. Castelli allowed copies of the letter to pass into circulation, and it fell into the wrong hands. In December 1614 a sermon was preached in Florence criticizing the opinion of Galileo for being contrary to Scripture and inimical to the Catholic faith, and soon a complaint was sent to the Holy Office in Rome. Early in 1615, Paolo Antonio Foscarini published a book, with admiring references to Galileo, showing how completely Scripture supports Copernican theory. He sent the book to Cardinal Robert Bellarmine,

Also Reviewed in This Issue

defender of the Church against Protestants and heretics. Bellarmine wrote a letter to Foscarini, also intended for Galileo, making these points: Although the Sun at the center and the Earth in motion may be taken as a supposition, to maintain it as true endangers faith by making Holy Scripture false, for Holy Scripture must be interpreted in accordance with the Fathers of the Church, all of whom agree that the Earth is immobile at the center of the world. Galileo drafted a reply to Bellarmine, arguing that the motions of the Earth are not just a supposition but can be proved, but he never completed or sent it.

By late 1615 there had been more complaints to the Holy Office. Galileo, as he told the Grand Duke, was stung by the vain slandering of his adversaries that he held erroneous opinions in his works, and resolved to come to Rome for the purpose of vindicating himself from such accusations and making clear the truth and his righteous and pious intention. He never should have gone; had he not done so, history would have been different. In Rome he did what he did best: He talked, for no one could talk more brilliantly, more convincingly. And what he talked of was Copernican theory, arguing, demonstrating, refuting, devising far more damaging objections to it than could anyone else, and refuting them too. But it did no good and probably did much harm.

Then Pope Paul V turned the matter over to Bellarmine, who asked eleven Father Theologians to the Holy Office to assess two propositions, the immobility of the Sun and the mobility of the Earth. On February 24, 1616, they ruled that the first was "foolish and absurd in philosophy and *formally heretical* as it expressly contradicts the teachings of Holy Scripture," and that the second "receives the same censure in philosophy, and with regard to theological truth is *at least erroneous in faith*." What began as a difference over interpretation of Scripture had become a condemnation of Copernican theory. On February 26, Galileo was warned to abandon entirely the opinion that the Sun is the center of the world and immobile and the Earth moves, and he was admonished to "*in the future not hold, teach or defend it in any way either by speech or writing*"; otherwise, the Holy Office would proceed against him. He acquiesced to this order and promised to obey.

That was the end of Galileo's intention to prove Copernican theory in his *System of the World*. He occupied himself with finding accurate periods of Jupiter's satellites to compute tables for determination of longitude, and, following the appearance of the bright comet of 1618, engaged in a controversy on comets and telescopes, and empiricism and authority in science, in *The Assayer*.

Then in 1623 Galileo's old friend and supporter from Florence, Cardinal Maffeo Barberini, was elected Pope Urban VIII. In audiences with the Pope the following year Galileo was given to understand that it was now permissible to write on the system of the world, provided that he do so hypothetically and stay away from Scripture. The Pope also repeated something he had said to him years earlier: Granting everything devised concerning the heavens and the motion of the Earth, does not God have the power or knowledge to arrange the spheres or stars in another way such that everything can be saved? And if God has the power and knowledge to arrange these in another way, we must not restrict the divine power and knowledge to this one way. Having heard this, Galileo was silent. This was more than a question about Galileo's science; it was a stern warning. And it makes all certainty in science impossible.

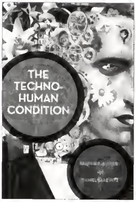
Galileo returned to Florence and began, or resumed, work on what became the *Dialogue on the Two Great Systems of the World, Ptolemaic and Copernican*, completed in 1630 and published in 1632. A more accurate title would be *Dialogue on the Two Great Systems of the World, Aristotelian and Galilean*, although Galileo, for all his nerve, knew enough not to say that. He did stay away from Scripture, but he did not treat the subject hypothetically, even though he made occasional remarks to that effect. What he wrote was a sustained argument of nearly 500 pages to prove Aristotle wrong and Copernicus—or, better, Galileo—right. He takes up, in order, these topics: circular motion and the nature of the heavens, which are subject to change; the diurnal rotation of the Earth, which it must have, and why it is not detected; the annual heliocentric motion of the Earth, which it must have and which can in fact be proved; and the cause of the tides from the diurnal and annual motions of the Earth, which proves that the Earth must have both motions. The *Dialogue* is brilliant and compelling: The arguments against Aristotle are devastating—any remaining Aristotelians must be fools—and the arguments for the motions of the Earth are effective, even though Galileo himself knew that some were flawed and Newtonian mechanics has shown others to be defective. He had accomplished just what he set out to do, prove Aristotle wrong and Copernicus right, and for this he expected celebrity and gratitude.

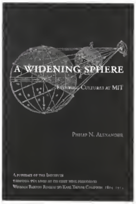
Galileo received both from his friends and former students and many throughout Europe, but he did not receive either from the person he was most interested in winning over. He believed that when his old friend Maffeo Barberini saw the strength of his arguments, saw the *truth*, he would rescind the prohibition against Copernicus and congratulate Galileo for proving the system of the world once and for all and for rescuing the Church from an unfortunate error. For if Galileo stood for anything, it was that in science, which is concerned only with truth, truth will prevail. But the Pope believed his own truth, that God has the power and wisdom to do things in ways no one can understand, and that to deny this challenges the omnipotence of God, which is dangerous to religion and faith. And Galileo had not only violated this truth but had treated it ironically and dismissively in the *Dialogue*. The Pope believed himself betrayed and insulted, and he appointed a commission to examine the *Dialogue*.

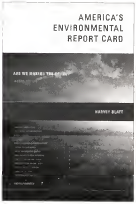
In September 1632 the commission presented a report reviewing the *Dialogue* and raising various objections. But it also found the record of the Holy Office from 1616 stating that Galileo had been warned to abandon entirely his opinion that the Sun is the center of the world and the Earth moves and "*in the future not hold, teach or defend it in any way either by speech or writing.*" Here was all that was needed to remove the matter to the Holy Office, over which the Pope himself presided.

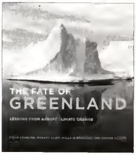
Galileo had four appearances before the Holy Office in Rome in April through June of 1633. In the first he insisted, against the evidence of the record, that he had received or could remember no such warning, and against the evidence of his book, that he neither held nor defended the mobility of the Earth and the stability of the Sun, but showed "the contrary of Copernicus's opinion and that Copernicus's reasons are invalid and inconclusive." This was, to say the least, hard to believe, and three subsequent reports concluded

The MIT Press









The Techno-Human Condition

Braden R. Allenby and Daniel Sarewitz

A provocative analysis of what it means to be human in an era of incomprehensible technological complexity and change.

256 pp., 1 illus., $27.95 cloth

A Widening Sphere

Evolving Cultures at MIT

Philip N. Alexander

"*A Widening Sphere* is a deeply researched and well-written account of the MIT's first nine presidents. It is also something more—engaging and even inspiring—unexpected but greatly appreciated qualities for an institutional history."

— Gerald Early, Washington University in St. Louis

432 pp., 45 illus., $29.95 cloth

America's Environmental Report Card

Are We Making the Grade?

Second Edition

Harvey Blatt

"This book serves as an indictment of inaction on the part of citizens and policymakers, a much-needed critique of the consumption-based economic drivers of environmental degradation, and a call to action for all Americans Blatt's revision of the first edition is a welcome and valuable contribution to the field."

—James W. Jordan, Antioch University New England

376 pp., 47 illus., $19.95 paper

The Fate of Greenland

Lessons from Abrupt Climate Change

Philip Conkling, Richard Alley, Wallace Broecker, and George Denton

photographs by Gary Comer

"This is the best accounting of abrupt climate change available—best because the science is on-target and nuanced and the storytelling is superb."

— Peter Kareiva, Chief Scientist and Vice President, The Nature Conservancy

240 pp., 78 illus., color throughout, $29.95 cloth

Quantum Computing

A Gentle Introduction

Eleanor Rieffel and Wolfgang Polak

"The authors have given us an introduction to the new field of quantum information, accessible to anyone familiar with college-level mathematics. It will be the easiest way for anyone to go from knowing no quantum mechanics to understanding cutting-edge problems in quantum computing. It will also be the most comprehensive and current book on the subject."

— Michael B. Heaney, Applied Quantum Technology Solar, Inc.

Scientific and Engineering Computation series
600 pp., 79 illus., $60 cloth

Transformations of Lamarckism

From Subtle Fluids to Molecular Biology

edited by Snait Gissis and Eva Jablonka

A reappraisal of Lamarckism—its historical impact and contemporary significance.

Vienna Series in Theoretical Biology
448 pp., 23 illus., $50 cloth

Chimeras and Consciousness

Evolution of the Sensory Self

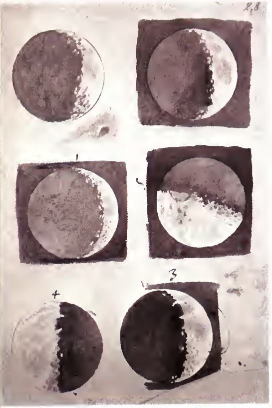
edited by Lynn Margulis, Celeste A. Asikainen, and Wolfgang E. Krumbein

"In this volume a group of bold and imaginative scholars probes the edges of the paradigm to investigate the 'hard problem' of consciousness by exploring its evolutionary roots from deep in the microbial world to its cultural embodiments. This is a new view of the biosphere, natural philosophy at its most challenging."

— Harold J. Morowitz, George Mason University

336 pp., 12 color illus., 42 b&w illus., $29 paper

To order call 800-405-1619 • **http://mitpress.mit.edu** • Visit our e-books store: **http://mitpress-ebooks.mit.edu**



Galileo made these wash drawings of the lunar landscape in 1609, probably with the help of an instrument like the early Galilean telescope pictured here. In instruments of this design, the field shrinks as the magnification increases; Galileo would have been able to see only about a quarter of the full moon at one time, so the drawings are composite portraits. This particular telescope was 92.7 centimeters long and was probably given as a gift to the Grand Duke of Tuscany, Cosimo II, in about 1610. From J. L. Heilbron's *Galileo*.

that Galileo held, taught and defended the Copernican opinion. After three further appearances, in which the most he would admit to was making the arguments for Copernicus—arguments he said he intended to refute—convincing rather than easy to answer, he repeatedly denied the charges against him. On June 22, in accordance with the order of the Pope, he was found "vehemently suspected of heresy," a formal term meaning in effect "seriously guilty of heresy." He was given the opportunity, really the necessity, to abjure his heresies, which he did, reading a humiliating statement written for him, as the alternative was imprisonment or worse.

His sentence included prohibition of the *Dialogue*, swearing that he would never again say or assert, in speech or in writing, things through which one could have similar suspicion of him (meaning the immobility of the Sun and the mobility of the Earth), and imprisonment at the pleasure of the Holy Office. In the following months, the sentence and abjuration were read to professors of philosophy and mathematics at universities throughout Catholic Europe, "so that they will understand the gravity of the error committed by Galileo in order to avoid it along with the punishment they would receive were they to fall into it." Galileo's imprisonment was later commuted to house arrest in his villa in Arcetri outside Florence, where he remained under close supervision that lasted the rest of his life, with repeated warnings that he speak with no one about the motions of the Earth.

But Galileo was not defeated, for under the severe conditions of confinement, he wrote the *Discourses and Mathematical Demonstrations Concerning Two New Sciences*, which was sent by way of Venice to Leiden and appeared in 1638. Here at last he published his discoveries in mechanics—some made many years earlier and now revised into their final form, others made very recently and worked out along with the revision. The mathematics, which is cumbersome and often obscure, was soon superseded in brevity and clarity by that of his successors—in truth Galileo was not a good mathematician. But the discussion in dialogue form is brilliant and filled with ingenious illustrations of the mechanics, paradoxes that continue to astound, and as devastating an attack on Aristotelian physics as anything in the *Dialogue*. By the time the book appeared, Galileo was blind, but that did not keep him from working. He kept up his correspondence through dictation, and in late 1641 began another dialogue eventually published as an addition to the *Two New Sciences*. He died on January 8, 1642.

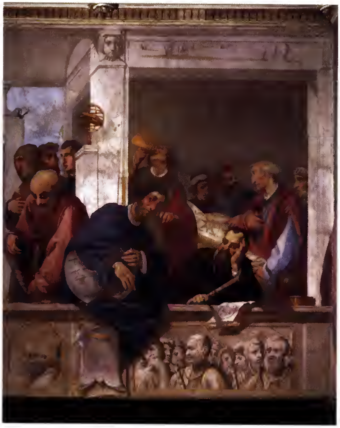
Galileo's lasting fame rests on his discoveries with the telescope, the reflections on science in *The Assayer*, the arguments for Copernicus and against Aristotle in the *Dialogue*, and the mechanics of the *Two New Sciences*, but most of all on his condemnation by the Church—the most famous event in the history of science, whose pertinence to contemporary issues never seems to fade. As long as there is a conflict, a contradiction, between science and religion, between reason and faith, between the natural and the supernatural—which appears inevitable from the essential character of each—Galileo will remain the most famous of all scientists.

The enormous literature on Galileo continues to increase at a rate faster than anyone can keep up with, certainly faster than I can. The foundation of modern studies is Antonio Favaro's monumental edition *Le Opere di Galileo Galilei*, published between 1890 and 1909; it is the source of all serious scholarship on Galileo, along with Favaro's many other publications. But these works are for scholars of Galileo. For everyone else, the two books I consider the most interesting and essential are Leonardo Olschki's *Galilei und seine Zeit* (1927) and Stillman Drake's *Galileo at Work* (1978). Olschki's is the third volume of a longer work on the practical mathematical and technical literature of the Renaissance, from which Galileo's own interests in mechanics arose and of which his own writing in the vernacular is a continuation. Drake is the foremost scholar of Galileo after Favaro, known for his translations of most of Galileo's important works and a lifetime of publication of books and articles on him. *Galileo at Work* is a life and work with the emphasis on the work as it took place, shown in letters and manuscripts.

The two most recent biographies, J. L. Heilbron's *Galileo* and David Wootton's *Galileo: Watcher of the Skies*, are both the product of immersion in Favaro's edition, from which, even after a century, new discoveries can be made. Both are also based on extensive research in primary and secondary literature. And both tell the story of Galileo's life in detail, with any number of incidents that will be new to nearly every reader. Both books, in my opinion, are fascinating, even difficult to put down. They are, however, quite different, and the principal difference concerns science.

J. L. Heilbron is a noted historian of science, primarily of physics, who has written on a broad range of subjects from the early modern period to the 20th century. In 1999 he published *The Sun in the Church: Cathedrals as Solar Observatories*,

a history of the research in astronomy encouraged and sponsored by the Church in the 17th and early 18th centuries. His new biography considers all the varied aspects of Galileo's life, including his interests in art, literature and poetry, beginning with his lectures in 1588 on the plan of Dante's *Inferno*. Heilbron also provides an account of Galileo's scientific work that is based on detailed knowledge and understanding of the science, which he explains very well. No one since Drake has treated as much of Galileo's scientific work. Because Drake's accounts are spread among numerous papers, collections and books, Heilbron's book, to the best of my knowledge, explains more of Galileo's science than any other single book. It is all covered here, from the early Archimedean theorems on the center of gravity of solids and writings on motion and machines, through the experimental and theoretical researches on accelerated motion and the parabolic trajectory of projectiles recorded in Galileo's manuscript notes, to the discoveries with the telescope, the hydrostatics of floating bodies, the controversies on sunspots and, later, comets and telescopes, to the arguments and demonstrations concerning the motions of the Earth in the *Dialogue* and the final explication of mechanics in the *Two New Sciences*. In places the exposition, with the mathematics, is compressed and will take some concentration to follow, but the many diagrams clarify the mathematics, and the full range of Galileo's science is explained so that anyone who studies it carefully can understand it. The conflicts with the Church in 1616 and 1633 are set out in appropriate detail and without the strange conspiracy theories that have grown up around this subject. Above all for its exposition of Galileo's science, Heilbron's book is worthy to take its place beside those of Olschki and Drake.



A number of illustrious Tuscans, including Galileo (holding a telescope), are depicted in this 1636 wall painting by Cecco Bravo, which Michelangelo Buonarroti the Younger commissioned to adorn his house in Florence. At the far left, the architect Antonio Manetti stands holding his model of Dante's Inferno, for which Galileo calculated the mathematics in 1588. From J. L. Heilbron's *Galileo*.

David Wootton is a historian, principally of political philosophy and of religious skepticism in the early modern period; most pertinent to the present volume is his 1983 book, *Paolo Sarpi: Between Renaissance and Enlightenment*. Sarpi, a Servite monk who became theologian to the Republic of Venice during the controversy over the Interdict of 1606, was a friend and correspondent of Galileo's during his years in Padua and later. If any contemporary was Galileo's equal, it was Sarpi, who was as radical as Galileo and as brilliant a writer. In his *History of the Council of Trent* (1619), Sarpi did to the Counter-Reformation Church what Galileo did to Aristotelian natural philosophy.

Galileo: Watcher of the Skies is more a biography than an account of Galileo's science. The book is divided into 39 sections, most fairly brief, which read like a series of essays on aspects of Galileo's life. There is nothing wrong with this, and many of the essays are lively and on subjects not often treated, such as Galileo's family and his relations with them—which consisted mostly, it appears, in giving them money. I wish, however, that Wootton knew more of the science and could explain it better. There is no real mathematics, and there are no diagrams, which precludes a proper exposition of Galileo's mechanics. And the treatment of astronomy is not better, with curious reflections on Galileo's "premature Copernicanism," which might not seem premature if the evidence that convinced Copernicus, Kepler and probably also Galileo were understood.

Wootton also falls for conspiracy theories of Galileo's troubles with the Church. In particular, he invents a scene in which he imagines Galileo's Inquisitor telling him that if he fails to cooperate in the proceedings against him, he will face an additional very serious charge, that of having denied transubstantiation—a charge based on the distinction he made between primary and secondary qualities in *The Assayer*. Wootton gives rather a lot of attention to arguing that Galileo was an unbeliever, as if this were something new. From all I know of Galileo and from his use of irony in his writing, particularly on anything to do with piety, it would surprise me if he were anything else. It would also surprise me if this were not obvious to the Holy Office when it compelled him to abjure with words that it was evident he could not possibly believe. And I wish that Wootton did not rail so much against writers with whom he disagrees, because no one who knows anything takes them seriously, and many of them are dead.

Still, Wootton has written a lively book that is interesting to read, and one can pass over the superfluous interpretations and concentrate on the fascinating details from the extensive research. And after all, it is about Galileo.

Noel M. Swerdlow, an emeritus professor in the department of astronomy and astrophysics and the department of history at the University of Chicago, is currently a visiting associate in history at the California Institute of Technology. He is the author of The Babylonian Theory of the Planets *(Princeton University Press, 1998) and is coeditor with Trevor H. Levere of a collection of 80 papers by Stillman Drake, titled* Essays on Galileo and the History and Philosophy of Science *(University of Toronto Press, 1999).*

Visualizing Disciplines, Transforming Boundaries

William J. Rankin

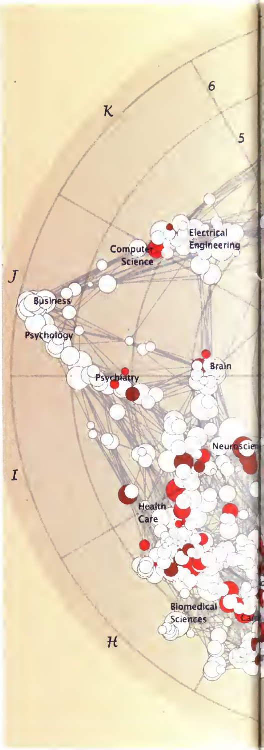
ATLAS OF SCIENCE: Visualizing What We Know. Katy Börner. xii + 254 pp. The MIT Press, 2010. $29.95.

Academic disciplines are a comfort and a cage: Their shared literature creates communities and defines common problems, but they can also inhibit the exploration of uncharted territory. Disciplinary boundaries frame the basic questions of research. What should I read? Where should I look for new ideas or collaborators? Just as important, when should I *stop* reading? These questions are hardly new, but of late they have clearly taken on a new urgency. Mass digitization has lowered the barriers to entering unfamiliar fields and made it easy to find common interests in unexpected places. Even going to the library is now a rare chore.

The premise of Katy Börner's *Atlas of Science* is that the inherent fluidity of scientific cross-pollination—in analog and digital forms alike—requires new tools for understanding scientific literature and disciplinary formations. Disciplines are not consistent and self-evident domains of knowledge, but rather transient, emergent phenomena that are reconfigured as quickly as they crystallize. The bulk of Börner's book is a detailed presentation of 18 "science maps" (that is, maps *of science*, rather than maps that are scientific) created by a diverse assortment of information specialists, engineers, scientists and designers. These are not geographic maps, but instead intellectual, social and conceptual ones; they mostly rely on the same bibliographic databases available to scholars, reoriented so that the tools of interdisciplinarity become evidence of their own impact. The maps were collected as part of an ever-evolving traveling exhibition curated by Börner since 2005, and they are now all freely available on the Web (at http://scimaps.org). The book thus participates fully in the trends it seeks to document. It is collaborative, cross-platform and interdisciplinary; it combines elements of the exhibition catalog, poster session, blog and monograph. Even Börner herself is a hybrid, with academic affiliations stretching from information science and statistics to cognitive science and biocomplexity.

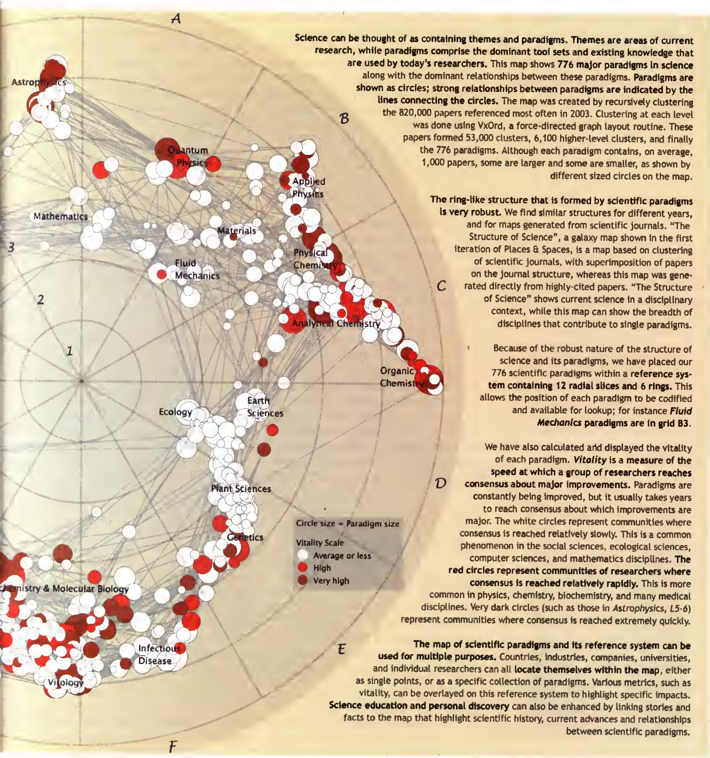
The *Atlas of Science* advances two simple propositions. One is that its various graphs and diagrams are indeed maps that make up an orderly atlas. The other is that the subject of these maps is in fact science, defined liberally and expansively. Both of these claims are more provocative than they at first appear, and together they make this book relevant well beyond the specialist world of information graphics.

Börner shows convincingly that mapping science and mapping the Earth are allied endeavors, but their unstated differences are obvious throughout the book. It is certainly true that maps need not be geographic—mapping is about relationships, not just topography. This large conceptual umbrella, however, hides an important methodological difference between the maps in this book and more familiar varieties. The success or failure of a traditional statistical map is measured by its ability to make data legible; it must present obvious patterns, highlight anomalies and eliminate ambiguity. The maps collected by Börner instead place much more emphasis on data processing than on clarity. The maps' descriptions tend to focus not on the authors' graphic choices but on the specifics of each database, the steps required to process it, and possibilities for future iterations. The difference here might be described as one between cartography (or information design) and visualization. Cartography has traditionally been concerned with trustworthy representation—digesting complex data and giving it fixed visual form. Visualization, in contrast, is usually an ongoing process of data exploration, in which any individual image always seems less important than the algorithms used to create it.



This map of scientific paradigms was created by Kevin W. Boyack and Richard Klavans in

This disconnect is both textual and visual. Börner spends her first three chapters establishing a long historical trajectory for science maps, complete with time lines, biographical vignettes and thumbnail reproductions of inspirational precedents. Yet the history of statistical cartography is surprisingly absent. Charles Joseph Minard's famous map-



2006 by clustering the 820,000 scientific papers referenced most often in 2003. As their accompanying commentary explains, the 776 paradigms identified are shown as circles, with lines between circles indicating strong relationships between paradigms. From *Atlas of Science.*

chart of Napoleon's retreat makes its requisite appearance, but that's about it. Nowhere do we find the groundbreaking census atlases of Francis Walker and Henry Gannett, the cartograms popularized by Erwin Raisz, or the "scientific cartography" of Max Eckert and Arthur H. Robinson. Jacques Bertin's powerful "graphic semiology" receives only a brief citation. The principal figures here are instead encyclopedists, librarians, computer scientists and futurologists—people such as Denis Diderot, Buckminster Fuller and J. C. R. Licklider. Börner's own disciplinary interest is clearly not in the practice of mapping, but in the organization, analysis and automatic processing of massive amounts of data.

The science maps themselves likewise defy the norms of traditional cartography. About half the maps are network graphs of various kinds; most of the rest are variations of word clouds and time lines. In all cases, the most compelling maps are those that promise more than could ever be delivered in print. Many trip over their own

vastness; they practically beg the reader to zoom in, click for more information or rerun the program with different parameters. The tangles of lines and color on the page might frustrate a reader seeking the clarity of a statistical atlas, but the dislocation between the evolving digital space of the algorithm and the finitude of the printed page is what makes these maps interesting. They point to new ways of understanding science precisely because they are *not* finished maps, but rather a new method of mapping altogether.

The version of science found in Börner's atlas is also a rather specific one, simultaneously expansive and focused. No domain of human knowledge is left out, and the unity of the sciences is taken as a given. The statistics marshaled here include everything from the contents of specialist physics journals to the growth of book production since the 17th century. Nevertheless, nearly all the maps share a consistent concern with the individual scientific publication as the unit of scientific practice. Some maps chart changes in narrow subspecialties, whereas others tackle all fields at once, but the focus remains on practices of authorship, especially complex networks of citations, collaborations and keywords. About half the maps rely on citation analysis alone, and all but three use databases of a similar sort: Wikipedia histories, patent activity, conference abstracts and so on.

This focus is largely to the book's credit. These sources are deep and rewarding, and the resulting maps show decisively that clear disciplinary boundaries are the exception rather than the rule. But it is relatively easy to imagine other ways that science might be mapped, especially once the idea of mapping includes things like word clouds. Could the methods used for tracking citations also generate maps of pedagogy, funding structures, conference attendance, lab techniques or professional organizations? Or even more broadly, what about maps of the effect of science on everyday life, the role of science in public policy, or the tensions between science and other ways of knowing? All of these are data-rich areas that could be well served by new forms of visualization.

Börner and her collaborators, however, never venture far from bibliography. Perhaps this means that the *Atlas of Science* simply falls short of its ambitious title. Perhaps it ought to have been called something like *Bibliographic Visualization* instead. But this kind of modesty would be unsatisfying, because the narrowness of the book's idea of science in fact ends up offering a rather radical vision of our disciplinary future, and its methods deserve the careful attention of scientists and designers alike. My own impression is that the book is more radical than even Börner would acknowledge. It is a kind of methodological wolf in sheep's clothing, with each synoptic, information-dense science map acting as a quiet manifesto against our slow, linear and bulky reliance on text.

At its most forceful, Börner's project is only partly about mapping the sciences. Instead it positions data visualization as an indispensable part of organizing and navigating scientific literature. This is a potentially monumental shift—a shift from words to images and from discrete legible units to the fuzzy impression of the overall pattern. It could easily rival the shift we have already seen from ordered indexes to search as the preferred mode of navigating text. The methodological disconnects between the science map and traditional notions of cartography (and science) should thus be taken very seriously, since science maps are not just a new graphic portrayal of scientific authorship. They are instead harbingers of new bibliographic methods that could lead to new ways of practicing and steering science itself.

What will happen when scientists see network graphs as a viable alternative to the literature review? Or when funding agencies start using visualization to encourage new disciplinary constellations? The vision of data-driven bibliography is an optimistic one. It offers fascinating new tools and proposes that we simply try them out. But this means that our job as researchers will be to ask how the tools can address the tasks that matter to us, rather than simply accepting what they can already do.

William J. Rankin is finishing a dual Ph.D. in history of science and architecture at Harvard University and will be an assistant professor of history at Yale University beginning in the fall of 2011. He is currently writing a book on the history of the mapping sciences, sovereignty and U.S. military globalism in the decades surrounding World War II, tentatively titled After the Map: Cartography, Navigation, and the Transformation of Territory in the Twentieth Century. *He also publishes maps of his own and maintains a website about mapping at http://radicalcartography.net.*

FORENSIC GENETICS

DNA Evidence

Simon A. Cole

THE DOUBLE HELIX AND THE LAW OF EVIDENCE. David H. Kaye. xviii + 330 pp. Harvard University Press, 2010. $45.

GENETIC JUSTICE: DNA Data Banks, Criminal Investigations, and Civil Liberties. Sheldon Krimsky and Tania Simoncelli. xviii + 406 pp. Columbia University Press, 2011. $29.95.

Over the past 25 years, DNA profiling has developed into an enormously powerful policing technology. This development, however, has raised a number of difficult legal, political and scientific questions. These include questions regarding such matters as how broad the government's ability to collect DNA samples should be, what sort of genetic information should be retained in government DNA databases, in what manner those databases should be searched, what degree of scientific acceptance is needed to warrant using a genetic analysis as forensic evidence, and how the results of genetic analyses should be presented to judges and jurors. These debates and others have spawned a number of new books, including *The Double Helix and the Law of Evidence*, by David H. Kaye, and *Genetic Justice*, by Sheldon Krimsky and Tania Simoncelli. Both books are valuable additions to the literature. They thoroughly cover two quite different aspects of the development of DNA profiling: The first examines the past, and the second reflects on the future.

The Double Helix and the Law of Evidence, although it touches on other is-

sues, is primarily concerned with the disputes over the legal admissibility of DNA evidence in the United States during the early and mid-1990s. That period saw fierce courtroom battles as government and defense attorneys litigated, first, whether DNA evidence was ready to be used in court, and, second, what sorts of statements about the value of the evidence expert witnesses were justified in making. Because DNA analysis was a new technology and because, unlike many other forensic techniques, it emanated from "high science," a number of renowned scientists, primarily molecular biologists, were mustered by both sides. As the courtroom struggles intensified, the debates between these heavyweights—scientists such as Eric Lander, Kenneth Kidd, Richard Lewontin and James Crow—bled over into scientific journals. The National Research Council (NRC) was soon drawn into the fray, convening two committees, which issued two somewhat contradictory reports in 1992 and 1996.

Author David H. Kaye is a law professor renowned for his meticulous attention to detail, careful argumentation and impressive technical mastery of statistical and scientific issues. He was also a player in the events described above: He consulted occasionally for defendants, served on the second NRC committee and published many scholarly articles commenting on the weaknesses of arguments made by courts, scientists and other legal scholars. *The Double Helix and the Law of Evidence* offers a detailed, authoritative accounting of the legal cases of this period and of scientific debates that ran in parallel in the pages of scientific journals.

The debates concerned a number of issues. One was the question of what constituted proper scientific controls over the interpretation of the output of analyses of genetic materials. Two questions regarding population genetics were also important: What can scientists assume about the degree of "structure" in the human population? And, is it justifiable to treat different genetic loci as independent when calculating the value of genetic evidence? Kaye offers a number of simple exercises and analogies to make this technical material accessible to lay readers.

Kaye tends to be more critical of defense attorneys and their experts than he is of the government, and occasionally he treats judicial insults of expert witnesses as statements of fact, rather than considering that they might be rhetorical ways of justifying the decisions judges want to make. But he is generally equitable in dispensing trenchant criticism, and few characters in the story (the statistical geneticist Newton Morton is one) escape completely unscathed. When Kaye inserts himself into the narrative, he strikes an appropriate balance, usually letting the reader know what he thinks about an issue without being at all self-aggrandizing.

Kaye notes at the outset that he is not a historian. This imposes more limitations than the trivial one he acknowledges—that he is not always accurate about his characters' precise ages. One limitation is that the book lacks context. For example, although Kaye is often critical of the quality of lawyering in the cases he discusses, only once does he mention the crippling lack of resources under which most criminal defense attorneys in the United States operate. (The O. J. Simpson trial, which Kaye analyzes in depth, may, with some justification, be viewed as the exception that proves the rule.) A reader looking for a better contextualized and more readable—although less legally and technically detailed—narrative of these events might turn to *Genetic Witness: Science, Law, and Controversy in the Making of DNA Profiling* (2007), by historian of science Jay D. Aronson.

Lack of context is also evident in the book's conclusion, when Kaye turns to the topic of lessons learned and discusses ways that the comprehension and use of science in legal proceedings might be improved. He focuses exclusively on the perceived biases of expert witnesses and the various mechanisms that have been proposed to mitigate those biases. Surprisingly, there is no discussion of whether biases (not to mention lack of scientific literacy) among judges might be an issue of equal, if not greater, importance. There is ample evidence in *The Double Helix and the Law of Evidence* that judges not only failed to understand the scientific issues well but may also have been reluctant to deprive the government of evidence that could be used to convict accused criminals. Although Kaye may view this as, in retrospect, the correct outcome, that does not necessarily mean it emanated from the right intentions.

Kaye concludes by declaring that "the great DNA wars over admissibility are over," but he acknowledges that thorny issues remain, such as so-called low-copy-number DNA analyses. In addition, although the issues relating to population genetics have been thoroughly aired, the question of how to interpret DNA analyses, especially analyses of "mixtures," samples that contain DNA from more than one individual, remains contentious and has received less attention from the courts, the NRC and scientific journals. Nonetheless, with the imprimatur of the courts, a public myth of the "infallibility" of DNA profiling has arisen.

Genetic Justice begins, in a sense, where Kaye's book leaves off—with this notion of infallibility. It recounts the notorious case of the "phantom of Heilbronn," in which German officials spent $18 million and 16,000 hours of overtime, analyzing 3,000 DNA samples taken without any direct suspicion, in the search for a supposed serial killer—a search based on a DNA profile that turned out to belong to a woman who worked at a factory packing swabs used for DNA sampling. Authors Sheldon Krimsky, a public-policy scholar well known for his work on the influence of corporate money on scientific research, and Tania Simoncelli, the former science advisor to the American Civil Liberties Union, note that it was the perceived infallibility of DNA profiling that prevented the authorities from considering alternative explanations for the results they were seeing.

In contrast to Kaye's book, *Genetic Justice* addresses the myriad privacy issues raised by the expansion of DNA profiling and DNA databases. Although DNA databases were initially limited to murderers and rapists, they rapidly expanded to encompass the perpetrators of other crimes and are now becoming an essential part of law-enforcement infrastructure. This expansion has been taken to its greatest extreme in the United Kingdom, where in 2008 the national database contained the DNA profiles of 4.6 million people (7.6 percent of the population). More than one-fifth of these profiles—about a million of them—were from individuals who had not been convicted of, or even charged with, a crime, but had merely been arrested. Although the all-encompassing nature of the U.K. database was recently deemed a contravention of the European Convention on Human Rights, the United States, among other countries, appears determined to follow the United Kingdom's example.

Genetic Justice offers a comprehensive discussion of contemporary civil-liberties intrusions associated with DNA profiling. In DNA "dragnets," DNA samples are demanded from inhabitants of a defined geographic area fitting a certain (usually rather vague) physical description. Krimsky and Simoncelli argue that these samples are not truly voluntary, because the police use intimidation and coercion to procure them. The authors also note that these dragnets are rarely effective at solving the crime in question and that the samples are rarely destroyed at the conclusion of the investigation; instead, they are added to the local DNA database.

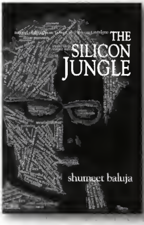
Familial searching is the practice of identifying suspects by searching the database for partial matches to a crime-scene sample. Inferring that individuals who match the sample at many but not all loci may be close blood relatives of the true source of the sample, the police generate suspects accordingly. As Krimsky and Simoncelli note, permitting such searches effectively adds to the DNA database the close blood relatives of convicts already in the database. *Phenotypic profiling* is the practice of attempting to predict the physical appearance of the perpetrator based on the genetic analyses of a crime scene sample. This practice raises thorny ethical and policy issues, in part because the technique sometimes purports to predict the perpetrator's race. Although such predictions are inherently probabilistic, this may not be well understood by users or by the public. Moreover, even predictions of phenotypic traits such as hair and eye color require inquiries into ancestry, because the frequencies of genetic markers associated with these traits vary across different ethnic groups. *Surreptitious sampling* occurs when law enforcement officers gather DNA samples from suspects without obtaining a warrant, simply by collecting ubiquitous "shed" DNA. Krimsky and Simoncelli note that under U.S. law there is no legal protection whatsoever for such "abandoned" samples, which bear the legal status of "waste." Moreover, courts have even upheld the use of samples obtained through deception, as in a famous case in which a suspect was tricked into providing a sample by licking an envelope. And, law enforcement aside, nothing prevents our fellow citizens from collecting our "abandoned" DNA.

Of course, not every country operates this way. Part II of the book puts the situation in the United States in perspective by surveying DNA databanking regimes in five other countries, a comparative approach also adopted by a recent volume edited by Richard Hindmarsh and Barbara Prainsack, *Genetic Suspects: Global Governance of Forensic DNA Profiling and Databasing* (2010). Krimsky and Simoncelli note that Germany, Belgium, Switzerland, the Netherlands and Norway mandate the immediate destruction of DNA samples after DNA profiles (which contain only a small amount of genetic information) are generated.

This policy of sample destruction emerges as perhaps the most appealing "technological fix" to civil-liberties concerns. Claims that current DNA profiles don't invade privacy because they are based on "junk" DNA that is completely devoid of predictive value have been overstated. Nonetheless, it seems clear that contemporary law-enforcement DNA profiles provide little basis for potential genetic discrimination. Thus destruction of the DNA samples used to generate the profiles would resolve many of the most pressing threats to privacy.

At the end of the book Krimsky and Simoncelli offer a reasonable set of additional policy recommendations: Apply the same protections afforded medical records to the information held in law-enforcement biological databanks; restrict the contents of the database to profiles of felons; and require warrants for dragnets, familial searches and recovery of shed DNA. Although these are reasonable suggestions, Krimsky and Simoncelli do not offer an entirely coherent ethical vision for how to grapple with the issues raised by the advent of DNA profiling. In particular, they tend to invoke all possible criticisms—invasion of medical privacy, race discrimination and ineffectiveness—without clearly articulating which of these issues is the fatal one. This results in some inconsistency: DNA databases are grave threats to privacy despite being ham-handedly ineffective; behavioral genetic research is both a pernicious resurgence of eugenics and a doomed research program based on a faulty un-

Princeton = Science



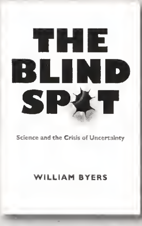
The Silicon Jungle

A Novel of Deception, Power, and Internet Intrigue

Shumeet Baluja

"At last, computer science has its equivalent to Scott Turow. Shumeet Baluja not only tells a compelling story, but as an expert in data mining, he also knows his stuff. His story shows how powerful and far-reaching modern search technology can be, and hence, potentially dangerous if not properly controlled. A techno-thriller with a disturbing message."
—Keith Devlin, author of *The Unfinished Game*

Cloth $27.95 978-0-691-14754-3



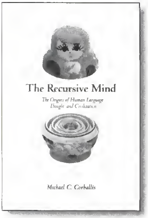
The Blind Spot

Science and the Crisis of Uncertainty

William Byers

"Byers has taken on a tremendously challenging task, one so daunting that it is hardly conceivable that it could be accomplished. *The Blind Spot* represents a serious advance, which in itself is very important and impressive."
—Reuben Hersh, National Book Award–winning coauthor of *The Mathematical Experience*

Cloth $24.95 978-0-691-14684-3



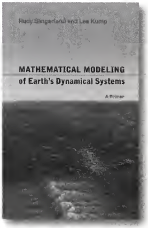
The Recursive Mind

The Origins of Human Language, Thought, and Civilization

Michael C. Corballis

"This is a wonderful book by an expert writer. Corballis tracks the importance of recursion in the context of language, theory of mind, and mental time travel, and concludes that its emergence explains much about how we became human. He proposes a novel answer to an enduring mystery. This book is a significant achievement."
—Thomas Suddendorf, University of Queensland

Cloth $29.95 978-0-691-14547-1



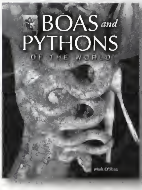
Mathematical Modeling of Earth's Dynamical Systems

A Primer

Rudy Slingerland & Lee Kump

"This wonderful, timely, and necessary book is a real winner. I appreciated the amazing range of geoscience topics as well as the book's structure—each of the chapters begins with an abstract-like summary preview, followed by examples of translations, before delving more deeply into topics. The authors should be congratulated for a brilliant book and pedagogical milestone."
—Gidon Eshel, Bard College

Paper $45.00 978-0-691-14514-3
Cloth $99.50 978-0-691-14513-6

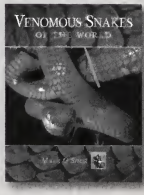


Boas and Pythons of the World

Mark O'Shea

"Well-known species and their more obscure cousins are all magnificently illustrated with beautiful color photos, with short write-ups of their life histories, range, size, prey, and other natural history. This excellent book is highly recommended."
—*Booklist*

Paper $19.95 978-0-691-15015-4



Venomous Snakes of the World

Mark O'Shea

"Poisonous snakes have been part of human legends—and nightmares—since the beginning of time. O'Shea, a reptile expert and television host, brings to life the world of serpents from around the word [with] hundreds of stunning, often rare, photographs."
—*Globe and Mail*

Paper $19.95 978-0-691-15023-9

PRINCETON UNIVERSITY PRESS

See our E-Books at
press.princeton.edu

derstanding of human differences. And, in my view, Krimsky and Simoncelli are too willing to credit genetic determinism in making their privacy arguments—too willing to perpetuate the view that "DNA sequence information may . . . contain information about behavioral traits, such as propensity to violence or substance addiction, criminal tendencies or sexual orientation." And they are too willing to credit a study that claims to explain marital discord genetically, perhaps in hopes that the claim will help raise the reader's hackles about genetic privacy.

Nonetheless, *Genetic Justice* constitutes the single most comprehensive articulation of the civil-liberties concerns associated with law-enforcement DNA databases and should, therefore, serve as a touchstone for debates about the spread of DNA profiling. David Kaye, meanwhile, is apparently working on a new book that will address privacy and civil-liberties concerns. Because Kaye is among the most vigorous defenders of the legality of government DNA sampling and storage, and is among the most persuasive advocates of a universal, population-wide DNA database, Krimsky and Simoncelli's arguments will soon find themselves with a worthy adversary.

Simon A. Cole is associate professor and chair of the department of criminology, law and society at the University of California, Irvine. He is a coauthor (with Michael Lynch, Ruth McNally and Kathleen Jordan) of Truth Machine: The Contentious History of DNA Fingerprinting *(University of Chicago Press, 2008) and is a contributor to* Genetic Suspects: Global Governance of Forensic DNA Profiling and Databasing, *edited by Richard Hindmarsh and Barbara Prainsack (Cambridge University Press, 2010).*

HISTORY

Of Passion and Polonium

Mary Jo Nye

RADIOACTIVE: Marie and Pierre Curie, A Tale of Love and Fallout. Lauren Redniss. 208 pp. It Books/HarperCollins, 2011. $29.99.

Lauren Redniss's *Radioactive* is a book that truly is out of the ordinary, and it is well worth reading and contemplating. Foremost, it is a work of visual art. That art is integrated with the text of a biography of Marie and Pierre Curie, which is interspersed with vignettes on the uses and perils of the radioactive elements that the Curies isolated and studied. Images appear on most pages, often alongside or as background for the text, which is set in a typeface created by Redniss. The typeface, which is not always easy to read, is based on the visual style of title pages of manuscripts in the New York Public Library and is named Eusapia LR, after the Italian Spiritualist medium Eusapia Palladino, who was popular in fin-de-siècle Paris and whom the Curies knew.

The book is illustrated in blues, oranges and reds, as well as in a lighter blue that is meant to suggest what Marie Curie referred to as the "spontaneous luminosity" of radium, which emits a faint-blue light. A cyanotype photographic printing process, which produces a rich Prussian blue background, is used for many of the author's drawings. This process, which uses sunlight and a solution of photosensitive chemicals, is meant to evoke similar techniques used to produce images in early radiation science. The bright orange and red colors in other images remind readers of the energy, the glow and the heat of radioactive substances; in one of the book's embedded vignettes, these colors are explicitly recalled by a survivor of the atomic bomb dropped on Hiroshima: "Suddenly all the windows in front of me became red. It was a beautiful color, like the sunrise mingled with orange."

In narrating the well-known story of Marie and Pierre Curie, Redniss admits that she has ignored advice from their granddaughter, Hélène Langevin-Joliot, to avoid telling the Curies' lives as a fairy tale. Redniss's narrative is not merely fanciful, however; it is documented with archival and published sources, including Eve Curie's biography of her mother, *Madame Curie* (originally published in 1937), and Susan Quinn's 1995 *Marie Curie: A Life*. Redniss's prose often is poetic, but some of the most striking poetic language is found in direct quotations of Marie Curie's writing, in particular Marie's 1923 biography of her husband, who died in a tragic street accident in 1906. Writing of their 1895 honeymoon in Brittany, Marie mentions that they loved "the reaches of heather and gorse, stretching to the very points of Finistère, which seemed like claws or teeth burying themselves in the water which forever rages at them."

After demonstrating the existence of two new and radioactive elements, polonium and radium, in 1898, the Curies laboriously succeeded in separating 0.1 gram of radium chloride in 1902. Marie described its "gleamings, which seemed suspended in darkness. . . . The glowing tubes looked like faint, fairy lights." In 1903 the Curies shared a Nobel Prize with Henri Becquerel, who had first discovered uranium's natural radioactivity.

Redniss's biographical narrative is spare but effective, concentrating on the relationship between Marie and Pierre and on their scientific work and their family; she includes the discovery of artificial radioactivity by their daughter Irène Curie and her husband Frédéric Joliot, and mentions the marriage in 1948 of Irène and Frédéric's daughter Hélène to Michel Langevin, the grandson of the man who had been Marie's lover decades earlier. Marie Curie's affair with Paul Langevin in 1910 and 1911 gets considerable attention (and artwork), as does Paul Langevin's duel with the gossip-mongering journalist Gustave Téry. Among the many striking images in the book is a hand-colored cyanotype print of a drawing of Marie Curie at the Nobel award banquet in Stockholm in December 1911, when she defied requests that she should not formally accept her second Nobel Prize because of scandalous publicity in the newspapers about her affair with Langevin.

As Redniss recounts, Marie Curie fully recovered her reputation following her volunteer work during the First World War, when she and Irène operated a fleet of mobile medical X-ray units at the front. Marie enjoyed triumphant tours in the United States in 1921 and 1929. In 1934, described by one of her assistants as so wraithlike that she appeared to be able to "walk through walls," Marie Curie died of anemia due to long radiation exposure.

The language and images of invisibility, immateriality, death and danger are the book's constant themes, in tension with visibility, matter, luminosity and love. The tension is captured early, in Redniss's dramatic remark that Wilhelm Röntgen's wife "intuited the power of her husband's discovery [of X-rays]—to intercept, as well as to hasten, death." Radium fascinated its beholders. The dancer Loïe Fuller asked the Curies to give her some radioactive material to illuminate the fabric of her veils as she whirled about the stage and the Curies' own home, but they refused.

Marie and Pierre recognized that radiation could have medical uses, and one of the vignettes that Redniss includes is a moving account of the use of radiation in 2001 to treat non-Hodgkin's lymphoma in an adolescent boy. Other vignettes are troubling or horrifying, including those that describe the many painful deaths that followed the bombing of Hiroshima in 1945, the effects on flora and fauna of fallout from the reactor accidents at Three Mile Island and Chernobyl, and the development of leukemia in children living downwind of the Nevada Test Site in the 1950s when atmospheric nuclear detonations were still being carried out there.

At first glance, *Radioactive* looks as though it might be a children's book, but it is instead a serious nonfiction work in the artistic and literary genre associated with the graphic novel and bande dessinée. An exhibit at the New York Public Library early in 2011 called "Radioactive" displayed 50 images from the book and other materials used in producing it; an interactive online exhibit (http://exhibitions.nypl.org/radioactive/) is still available.

Mary Jo Nye is Professor of History Emerita at Oregon State University. She is the author of Blackett: Physics, War, and Politics in the Twentieth Century *(Harvard University Press, 2004) and* Michael Polanyi and his Generation: Origins of the Social Construction of Science *(to be published by the University of Chicago Press in fall 2011).*

Marie Curie won her second Nobel Prize in 1911, in chemistry. As the text on these two pages from Lauren Redniss's *Radioactive* explains, after the prize ceremony in Stockholm, there was an eleven-course dinner at the royal palace, and Curie sat across from King Gustaf V. The Nobel Committee, worried that Curie's "stained reputation" (her love affair with Paul Langevin had just become public) would offend the royal family, had hoped she wouldn't attend, but the dinner went smoothly. Redniss notes that "before King Gustaf's reign was over, he, too, would be accused of a love affair, also with a married man." The musical entertainment at the dinner included "a piece from Georges Bizet's *Carmen* (in which a temptress inspires duels before being stabbed to death by her spurned lover) and the intermezzo from Moret's *Cleopatra* (a seductress's fate is sealed by a cobra bite)."

HISTORY

Finding Meaning in the Martian Landscape

David DeVorkin

GEOGRAPHIES OF MARS: Seeing and Knowing the Red Planet. K. Maria D. Lane. xiv + 266 pp. University of Chicago Press, 2011. $45.

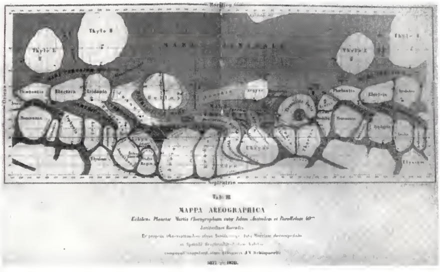
What can yet another history teach us about the canals of Mars? In the case of K. Maria D. Lane's new book, *Geographies of Mars*, the answer is,"a surprising amount"—about ourselves, our institutions and our authorities, and about what constitutes evidence and argument; the list is very long. The volume is an exceptionally well-written and cleverly crafted exposition of what both speculative and mainstream science had to say in the late 19th and early 20th centuries about the nature of Mars and the beings that might inhabit it; Lane situates this history in the broad cultural landscape in which the emerging professions of both astronomy and geography were being planted, tested and nurtured.

Lane's book appears at a time when there has been a welcome resurgence of scholarship recasting and extending humanistic explorations of the debate over life on Mars that took place in that era. *Geographies of Mars* provides insights that help the reader appreciate the resurgence. Lane avoids specialist jargon, making this literature accessible to readers whose background is in the sciences rather than the humanities.

In six taut and closely reasoned chapters, the author, an assistant professor of geography at the University of New Mexico, introduces the historical debate and the literature that has examined it from various perspectives. The idea that there were canals on Mars began with Giovanni Schiaparelli, who followed Mars through the weeks of its perihelic opposition in 1877 and for several months afterward, making his observations from a rooftop cupola in Milan with an 8-inch Merz refractor. In 1878 he published a memoir of his observations, which included a detailed map of the northern hemisphere of Mars showing islands and peninsulas divided by narrow blue bands of water that he labeled *canali* (Italian for *channels*); these soon became known as canals in the English-speaking world. Although no one confirmed his observations until 1886, the map was hugely influential, for reasons that include the regard in which Schiaparelli was held and the perceived objectivity of cartography. Updated maps of the canals that Schiaparelli produced in the 1880s became more and more abstract and geometrical. In 1895, the American amateur astronomer Percival Lowell argued that the "unnatural" and "artificial" appearance of the canals suggested that there was intelligent life on Mars, and he began producing his own detailed maps showing even more canals. He wrote books, went on lecture tours and spoke out in the popular press, making many claims about the landscape and civilization of the planet—that Mars was growing increasingly arid as a result of planetary evolution (and that a similar future awaited Earth), and that this had led the inhabitants to unite in peaceful cooperation to ensure a water supply. Many professional astronomers were quite critical of Lowell, but it wasn't until 1907, when photographs of Mars failed to show the canals, that his maps began to lose their authority. By 1910 most astronomers had concluded that the supposed canals were an optical illusion.

The combination of conditions under which Schiaparelli made his observations has always made me skeptical of what it was he was seeing, and there are no new scientific explanations in Lane's book. But by the time I finished reading her third and fourth chapters, I better understood why Schiaparelli's work was so appealing, and indeed why followers of his such as Lowell were as effective as they were in keeping the idea alive that a morally, socially and technically advanced civilization existed on the planet—a populace that, over millennia, had been driven by necessity to evolve into beings capable of reorganizing the geography of their dying home at a global level in order to survive and, evidently, to prosper.

Of course, those beings I speak of as inhabiting Mars are humans. They are us. Lane's book has convinced me of this, satisfying the suspicions first raised in me by pioneering historical works by William Graves Hoyt, Steven J. Dick and Michael J. Crowe in past years, more recently followed by David Strauss's excellent 2001 cultural biography of Percival Lowell. My suspicions were greatly heightened by Robert Markley's recent *Dying Planet: Mars in Science and the Imagination* (2005), a study that neatly juxtaposes fact, speculation and popular fiction. One can find this idea in all these works, more or less. But the realization didn't hit me until I read Lane's explanation of why the popular image of a so-



Italian astronomer Giovanni Schiaparelli included a Mercator projection map of Mars in his first Mars observation memoir, published in 1878. In a dramatic departure from previous maps, the northern hemisphere of Mars, shown in the bottom half of the map, is depicted as a landscape divided by many narrow waterways, which Schiaparelli labeled *canali* (Italian for *channels*). The darker portions of the map were originally printed using a bluish-green tint. From *Geographies of Mars*.

cially and technologically advanced civilization on Mars was distinctly, though not exclusively, an American cultural phenomenon. She relates the very existence of that image to a deep-rooted but definitely flawed belief among Americans that has been described as *exceptionalism*. In the conclusion to her book, she quotes cultural geographer Fraser Macdonald's observation that the exploration of space, "from its earliest origins to the present day, has been about familiar terrestrial and ideological struggles here on Earth." One can see clearly from the scope of Lane's work how contingent the appeal of advanced life on Mars must have been upon the American popular self-image.

Lane's physical and human geographical exploration plows through both the scientific investigations and the cultural reactions. For many readers this will be familiar ground. But it is ground tilled in a very different pattern, allowing for a broader interpretive cut. By employing both descriptive and explanatory arguments from cultural or human geography, and by showing how these elements influence not only the production of knowledge but also its dissemination and the subsequent assignment of authority, she manages to put into better focus issues that have been acknowledged in more traditional historical treatments (including my own), so that they can now be appreciated for what they are, not just within professional communities but in culture writ large.

Let me offer an example of how the book was helpful to me. I did a brief study in the 1970s, well summarized in Lane's book, regarding the spectroscopic search for water on Mars in the early 20th century, which I took to be the most sensitive indicator of the state of the art of scientific practice in that day, more objective than squinting at canals. I highlighted an episode in which astronomer William Wallace Campbell took his spectrographs to the highest and driest peak in the Rockies to rid his data of as much earthly contamination as possible. In my mind, the fact that the observations were conducted at high altitude was enough to lend authority to them, and I did not belabor the obvious. However, I never considered the possibility that their authority also arose from the extreme effort itself—the fact that they were made after an expeditionary effort to an exceedingly remote and hostile place. The expedition was a defining career milestone for Campbell, even more so than those on which he chased solar eclipses around the globe. In the larger context that Lane provides, I can better appreciate the depth of Campbell's frustration when his authority was called into question by negative popular reaction. I can also better appreciate why astronomers were so deeply frustrated by Lowell's challenge of their hegemony, a challenge that went far beyond the issue of site superiority to the question of the legitimacy of professional scientific practice.

Readers of *Geographies of Mars* who begin the book feeling satisfied with purely psychological or physiological explanations for why observers found canals on Mars in the Schiaparellian and Lowellian era will come away with an enlightened appreciation of how complex the problem really must have been. The book is a must-read for any historian or scientist who cares about what, how and why, and to what extent, cultural forces shape both scientific knowledge and public reaction to it.

David H. DeVorkin is Senior Curator, History of Astronomy and the Space Sciences, at the National Air and Space Museum of the Smithsonian Institution. He contributes to historical scholarship in his areas of curation and recently built a free public observatory on the National Mall in Washington, D.C.

PHYSICS

Feynman's Legacy

Silvan S. Schweber

QUANTUM MAN: Richard Feynman's Life in Science. Lawrence M. Krauss. xviii + 350 pp. W. W. Norton and Company, 2011. $24.95.

Richard Feynman (1918–1988) was one of the greatest physicists of the 20th century. His most original contribution, the path-integral formulation of quantum mechanics, has turned out to be hugely generative. In the nonrelativistic domain, Feynman conceived a microscopic entity (such as an electron) going from space-time point (x_1,t_1) to space-time point (x_2,t_2) as being able to take *any path* joining these two points, and he assigned a complex amplitude to each possible path that depended on the force field present and on the particular path. All these amplitudes were to be added, and the absolute square of that sum yielded the probability that the particle would go from (x_1,t_1) to (x_2,t_2). Feynman was able to determine which particular assignment of weight for each path yielded the usual Schrödinger formulation of quantum mechanics. In Feynman's formulation, a "particle" is clearly something very different from its classical conceptualization.

To generalize the theory to describe quantum mechanically the interaction of charged particles with the electromagnetic field, Feynman introduced a graphic shorthand to help him translate all the terms in the perturbative expansions of his integral over paths into calculable expressions for the transition amplitudes being considered. Not only did these diagrams eventually come to represent all the contributions to a given process, such as an electron being scattered by a proton and emitting a photon, they also became an elegant way to organize and *visualize* the perturbation series. Each diagram came to stand for the algebraic elements that went into the integral associated with its contribution to the perturbation series. As Peter Galison has stressed, the diagrams made calculations of any given process in Feynman's scheme a *modular* matter. One would draw all the diagrams that contribute to the process to the given order of perturbation theory one was to calculate, invoke the rules that associate particular components of the diagrams with mathematical expressions, and then calculate the integrals. Few elements of theory have become as omnipresent and as pervasive as Feynman diagrams.

Using this approach, in 1948 Feynman showed how quantum electrodynamics (QED) could be made to yield finite results for all observable quantities in the lowest orders of perturbation theory. Julian Schwinger and Sin-Itiro Tomonaga had demonstrated the same thing, and the three men shared the Nobel Prize in Physics in 1965 for their work on QED. In 1953 Feynman formulated a fundamental theory of liquid helium justifying the earlier phenomenological theories of László Tisza and Lev Landau.

In 1957, after Tsung-Dao Lee and Chen-Ning "Frank" Yang had estab-

lished that parity was not conserved in the weak interactions, and Richard Garwin, Leon Lederman, Valentine Telegdi, Chien Shiung Wu and others had experimentally corroborated Lee and Yang's finding, Feynman found he could account for all the experimental results by postulating that only the "left-handed" component of the wave functions of the particles involved in the weak interactions enter into the description. With Murray Gell-Mann, Feynman wrote an important paper on the subject that stimulated a great deal of subsequent theoretical activity in high-energy physics.

In the 1961–62 and 1962–63 academic years, Feynman taught the two-year introductory physics course that all Caltech students take. The lectures that he delivered were subsequently published in three red-bound volumes titled *The Feynman Lectures on Physics*. They have deeply influenced subsequent generations of physics teachers and physics graduate students.

When in the late 1960s experiments on the scattering of high-energy electrons by protons at the Stanford Linear Accelerator yielded very large cross-sections for inelastic scattering, Feynman came to the conclusion that the proton contained subnuclear entities he called *partons*, and that the experimental results could be interpreted as elastic scatterings of the electrons by these partons. Partons soon became identified with Gell-Mann's quarks, which are basic components of quantum chromodynamics.

Feynman's participation in the making of atomic bombs at Los Alamos during World War II generated in him a lifelong deep interest in computing and computers. His lectures on computing during the early 1980s, the series of papers on computing that he wrote between 1981 and 1985, and his contributions to the building of Danny Hillis's massively parallel computer mark the beginning of a new era of computers based on atomic elements operating by quantum logic.

Feynman achieved national prominence with the publication in 1985 and 1988 of *"Surely You're Joking, Mr. Feynman!"* and *What Do YOU Care What Other People Think?*, two books of stories from his life as told to Ralph Leighton. Feynman was even more in the national limelight in 1986 when, as a member of the presidentially appointed committee to investigate the space shuttle *Challenger* disaster, he dramatically pointed out its likely cause by dropping a rubber O-ring into a glass of ice water.

Feynman has been the subject of several biographies. James Gleick's splendid *Genius* (1992) sensitively narrated the physicist's personal life and made accessible to the general public many of his scientific contributions, particularly his work on QED. Jagdish Mehra's *The Beat of a Different Drum* (1994) is much more demanding technically. And now Lawrence Krauss has written *Quantum Man: Richard Feynman's Life in Science*, which is the latest entry in Norton's Great Discoveries series. In the introduction Krauss states that he was asked to produce "a short and accessible volume that might reflect Feynman the man as seen through his scientific contributions." He adds that his goal in writing the biography was to "reveal to nonphysicists . . . why Feynman has reached the status of a *mythic hero* [emphasis added] to most physicists now alive on the planet."

The book has much to recommend it. Krauss rightly stresses that Feynman always had his own way of looking at problems and his own way of solving them, and he notes that Feynman might have been able to contribute more had he not insisted on solving problems in his own way. Krauss's presentation of a good deal of Feynman's physics is insightful and should be accessible to nonphysicists. He is very good when narrating those parts of Feynman's physics he has had contact with—QED, weak interactions and quantum computing—but less so when discussing areas more distant from his own researches. I believe that his presentation of Feynman's explanation of the properties of liquid helium will leave the reader baffled.

I have other reservations as well. I do not believe that a book that is intended to convey the import and meaning of Feynman's scientific contributions to the public at large, and in particular to inspire a new generation of young people to enter the scientific enterprise, should underpin its presentation by reinforcing the mythic element. Krauss attributes the origin of many of the great advances in theoretical physics during the second half of the 20th century to some work that Feynman did. Because his emphasis is on the individual, the role of the community is lost, and the importance of the contributions of other remarkable individuals gets minimized.

Thus in Krauss's history of the postwar advances in QED, Freeman Dyson's contributions, although lauded, are not fully appreciated, for it is made to appear that all Dyson did was rederive Feynman's approach from the more conventional formulation of QED. (Krauss refers to Dyson having "helped the rest of the world understand QED, while establishing Feynman's methods as the ones that would ultimately root and grow.") What Krauss does not convey is that Feynman's formulation was based on a particulate conception of the basic entities. The antisymmetry of wave functions or propagators for electrons, and their symmetry for bosons (as in the case of liquid helium), had to be put in by hand. Nor is anything said about Feynman's inability to incorporate spin into his path-integral formulation, or about the fact that, because of this, the propagator approach described in his epochal 1949 *Physical Review* papers is not based on his integral-over-paths. One of Dyson's important contributions was to formulate Feynman's approach and its all-important diagrammatic component as a *field theory*, with the statistics of the "particles" the theory describes incorporated *ab initio* by virtue of the commutation rules obeyed by the field operators. Another of Dyson's contributions was to prove the renormalizability of the S-matrix that describes scattering processes in QED to *all* orders of perturbation theory, and to formulate the renormalizability criteria for quantum field theories. The importance of Feynman's path-integral approach in gauge theories is a consequence of its generalization—by others—to field systems, its treatment of Fermionic fields as Grassman variables.

Reading *Quantum Man*, one gets the strong impression that every subsequent development in QED is a consequence of "Feynman's QED." However, contrary to what Krauss suggests, the seminal contributions of Gell-Mann and Francis Low when investigating the small-distance behavior of QED were not based on "Feynman's QED," but on advances—by Schwinger, John Ward, Gunnar Källen and others—that occurred between 1949 and 1954; these advances had deepened the meaning of renormalizability and had formulated QED in the Heisenberg representation. There is in fact no reference to Feynman in the paper by Gell-Mann and Low.

There are other problems too. Krauss asserts incorrectly that Fermi got the Nobel Prize for his theory of beta decay. He states that the attraction for Feynman in beginning to work on quantum gravity in 1960 was that, as Feynman put it, "few of the best men are doing

work in [the field]"; although Krauss does then acknowledge that Feynman's remark "was probably somewhat of an overstatement," he makes no mention of the work on the quantization of general relativity that Bryce de Witt, Schwinger, Richard Arnowitt, Stanley Deser, Charles Misner, John Wheeler and others had been doing since the mid-1950s. Nor was Feynman the first to consider gravitational interactions as mediated by the exchange of spin-2 massless quanta. Not until late in the book do we hear of the role that Chen Ning Yang and Robert Mills's gauge theory (and Ryoyu Utiyama's insights into gauge principles) played in Feynman's approach. And absent is any talk about Peter W. Higgs and broken symmetry in gauge theories.

But what I find perhaps even more distressing is the absence of any discussion of the contexts that shaped and solidified Feynman's approach to physics and its theories. Peter Galison's highly informative and very insightful 1998 article, "Feynman's War: Modelling Weapons, Modelling Nature," links the modular theoretical culture of wartime Los Alamos to Feynman's modular approach to theory. This is the kind of insight that I believe must be conveyed to the public at large in order for people to understand the workings of science and the growth of the knowledge its practitioners produce. Similarly, one would like to have a more detailed assessment of what is characteristic of Feynman's *Lectures on Physics*.

Furthermore, perhaps in an attempt to humanize the "mythic hero," Krauss periodically informs the reader of Feynman's womanizing proclivities after the death of Arline, his first wife. It would have been more appropriate to put greater stress on how unusual a human being Feynman was in the way that he cared for the dying Arline when he was still a very young man, and to make the reader understand why this presumably highly rational human being wrote Arline a tear-drenched letter 16 months after her death, telling her of his love for her but not knowing where to send it.

If I have been critical of Krauss's book, it is because my hopes for it were quite high. As Feynman wrote to the German editor of *"Surely You're Joking, Mr. Feynman!,"* negative reviews are written "by people who expected more and were disappointed."

Silvan S. Schweber is professor of physics and Richard Koret Professor in the History of Ideas, emeritus, at Brandeis University. He is the author of a number of books, including QED and the Men Who Made It *(Princeton University Press, 1994) and* Einstein and Oppenheimer: The Meaning of Genius *(Harvard University Press, 2008).*

METEOROLOGY

From Rainmaking to Geoengineering

Rasmus E. Benestad

FIXING THE SKY: The Checkered History of Weather and Climate Control. James Rodger Fleming. xvi + 325 pp. Columbia University Press, 2010. $27.95.

When I was studying cloud microphysics in Socorro, New Mexico, back in 1992, I sometimes heard anecdotes about rainmaking and weather modification. However, I never realized how extensive the history of such efforts is until I read James Rodger Fleming's *Fixing the Sky: The Checkered History of Weather and Climate Control*. Fleming is a professor of science, technology and society at Colby College and has an academic background in astronomy, atmospheric science and history. The book provides a detailed account of weather modification, discussing the role it has played in mythology, literature and science fiction, and describing its history, covering the rainmaking efforts of both scientists and con artists, as well as such topics as fog removal in the early years of aviation and the weaponization of weather control. Fleming also gives an account of weather modification's standing today in a concluding chapter about current ideas and proposals that aim to address climate change through geoengineering.

When I started reading the book, I was a bit put off by the rather disjointed and sometimes tedious first chapter, which discusses how the theme of weather modification has been treated in Greek mythology and in a wide range of works of art, from Dante's *Inferno* to Disney cartoons. But after having read the rest of the book, I became more appreciative of the first chapter's discursions. Because the tales related there deal with attempts to control nature, they raise ethical questions and highlight the dangers of hubris, and that gives them relevance to current policy debates about whether to pursue geoengineering solutions for anthropogenic global warming.

Fleming makes it clear right away that he is no fan of weather modification, and he critiques it from a number of different angles. Charlatans and folly are common in his stories, and some of the historical episodes he describes are quite absurd in hindsight. There are some good stories—many past efforts

Humankind has a long history of confronting storms with a display of military might. In this woodcut from Olaus Magnus's *Historia de Gentibus Septentrionalibus* (1555), medieval archers attempt to deter a hailstorm. In later centuries, cannons were fired at storm clouds in the belief that this would disrupt the formation of hail or dissipate a storm's energy. From *Fixing the Sky*.

to control the weather appear to have been worthy of an Ig Nobel Prize. Fleming says in the preface that he views this history as a tragicomedy—indeed, a farce—and he can be quite sarcastic and ironic in his treatment of it.

In general Fleming gets the science right, with only minor inaccuracies and an occasional lack of precision. The book is not too technical and includes some nice illustrations. Most people will find it accessible. However, some passages are a little confusing on first reading, and it is not easy to get a clear overview when confronted with such a gigantic mosaic of anecdotes, so varied in chronology and theme. Nevertheless, the tales are intriguing.

The book's last three chapters are probably the most interesting. "Weather Warriors" looks at links between scientists and the military and between meteorology and the military, and at instances of weather modification that have had a military connection. During the Cold War, there was talk of seeding clouds to produce droughts or flooding that would hamper the enemy, and of creating hurricanes that could cause as much damage as a nuclear weapon. And during the Vietnam War, the U.S. Air Force seeded clouds with the intention of producing rainfall that would make travel along the Ho Chi Minh Trail more difficult.

The chapters titled "Fears, Fantasies, and Possibilities of Control" and "The Climate Engineers" are about proposals for geoengineering the global climate. Fleming portrays geoengineers as being like the weather warriors in that both groups appear to be attracted to climate modification as a macho activity that shores up their masculinity—as boys who want big, shiny, expensive toys to play with. But this culture is not the only problem.

One of Fleming's main arguments against geoengineering is that climate is unpredictable. However, the very existence of climate models means that we believe that there are aspects of climate that are to some extent predictable. The question of predictability depends on spatial and temporal scales and the degree of precision aimed for. Climate models can realistically reproduce large-scale patterns of temperature, precipitation and circulation. But when it comes to the prediction of regional patterns of temperature and rainfall, Fleming makes a good point—I think it is legitimate to say that we will not be able to predict the outcome of geoengineering for a specific country.

As Fleming emphasizes, geoengineering raises ethical and legal issues. Who has the right to control the climate? We know that even adaptation measures such as irrigation and dam construction can cause conflicts. That anthropogenic global warming is such a hot-button political issue is of course a bad omen for achieving any kind of consensus about what sort of geoengineering—if any—to undertake. And the potential for international conflicts seems enormous. Furthermore, the implementation of a geoengineering project on a global scale would require great sums of money; imagine lobbyists trying to influence decisions about how those tax dollars would be spent!

The topic is an important one, and the book is relevant for scientists, stakeholders, policy makers and concerned citizens alike. I already had my doubts about the feasibility of geoengineering before I read the book, but now I am convinced that it will open a can of worms.

Rasmus E. Benestad, a physicist, is senior scientist at the Norwegian Meteorological Institute and a contributor to the RealClimate blog (www.realclimate.org).

GENETICS

Making Sense of the Genomic Revolution

Simone Vernez and Sandra Soo-Jin Lee

THE LANGUAGE OF LIFE: DNA and the Revolution in Personalized Medicine. Francis S. Collins. xxviii + 332 pp. HarperCollins, 2010. $26.99.

THE $1,000 GENOME: The Revolution in DNA Sequencing and the New Era of Personalized Medicine. Kevin Davies. x + 340 pp. Free Press, 2010. $26.

HERE IS A HUMAN BEING: At the Dawn of Personal Genomics. Misha Angrist. x + 341 pp. Harper, 2010. $26.99.

For more than a decade, the public has been primed to expect an imminent revolution in genomics that would be the key to overhauling a broken medical system. Indeed, the promise of personal genomics is tantalizing: It could improve individual health by providing personalized risk information about diseases such as cancer, diabetes, heart disease and obesity, and about how individuals metabolize drugs, whether they are carriers for certain diseases and even what personality traits they are likely to have. This promise suggests a different landscape for health care, one in which individuals would actively participate in understanding and shaping their personalized health profiles. Equipped with their personal genomes, patients could choose healthier lifestyles and thereby achieve better health outcomes. Still, the promise of personalized medicine must be weighed against the challenges posed by the technological, financial, ethical and cultural limitations to which it is subject.

Three recent books—*The Language of Life: DNA and the Revolution in Personalized Medicine*, by Francis Collins, *The $1000 Genome: The Revolution in DNA Sequencing and the New Era of Personalized Medicine*, by Kevin Davies, and *Here is a Human Being: At the Dawn of Personal Genomics*, by Misha Angrist—provide accounts of the developments in personal genomics. Each of these authors offers unique insights into the possible benefits and costs of genomics research and its use in personalized medicine. The books also provide glimpses into how genomics may reshape our ideas about what constitutes health and well-being.

Collins, former director of the Human Genome Project at the National Human Genome Research Institute and current director of the National Institutes of Health, stands at the forefront of the genomic revolution he describes in *The Language of Life*. His compelling and accessible account details the potential of genomics to transform health decision making for individuals, predict disease risk, improve prenatal and postnatal health care, and expand our understanding of the vast microbiome with which we coexist. Collins recognizes that two obstacles facing personalized medicine are the absence of an adequate infrastructure and the lack of educational opportunities that support the translation of discoveries at the bench into useful applications at the bedside.

Acknowledging that clinical use of genome information is still in its nascency, he emphasizes that we must continue to invest in research, computerize medical record keeping, make appropriate policies, educate communities, and consider the ethical responsibilities confronting researchers and medical professionals in the face of changing technology.

Collins suggests that the key to realizing the potential of genomics is the active involvement of stakeholders, including scientists, doctors, policy makers and individuals. To this end he offers his book as a guide. For example, he urges that people consider three factors in deciding whether or not they want to know about their risk for a given illness—the baseline risk for the disease, the disease burden and the interventions available—and he emphasizes the importance of seeking knowledge that will equip them to make better decisions about prevention. In addition, at the end of each chapter he lists resources and activities that give readers the chance to "join the personalized medicine revolution." Collins writes that it's not enough to follow the progress of the revolution; we must "skate where the puck is going to be." The message is that individuals must assume an active role in harnessing and even shaping the potential of genetic medicine. However, it is unclear how early discoveries will translate into meaningful differences and whether our current infrastructure will support the incorporation of genomic data.

The price of having one's genome sequenced has been dropping rapidly. The first genome sequenced by the Human Genome Project cost an estimated $3 billion and took 13 years to accomplish, but by 2009 the task could be completed in less than two weeks at a cost of $1,500. It has long been asserted that the $1,000 genome would be the gateway to a revolution in health care, allowing a new medical and scientific reality to be realized. Critically acclaimed science writer Kevin Davies has a long track record of reporting on human genome sequencing; his new book, *The $1,000 Genome,* offers an intimate history of the technology and people behind these developments. Davies introduces readers to the key actors who have led the way in the major developments in genetic technology, revealing the unabashed optimism of the scientists and entrepreneurs who are trying to bring genomic sequencing ever more quickly and cheaply to the public.

Davies describes the new obstacles that are arising as the industry approaches the $1,000 threshold. Scientists must now look beyond sequencing technology and financial considerations, and consider the wide gap between the abundance of the newly available data and the paucity of knowledge about how to interpret it. Limitations of the technology—mainly how little is yet known about how to successfully translate genetic information into generalizable and reliable clinical information—pose significant barriers to the achievement of the promise of personal genomics. In addition, social and ethical issues, such as the specter of eugenics and risks for breaches of privacy and racial stigmatization to occur, have yet to be addressed. A reliable framework for guiding the successful application of genetics in research and clinical medicine is needed. Davies provides a historical account, from insiders' perspectives, of the evolution of genetic sequencing technology and the challenges that has entailed, painting a detailed picture of what we as a society must consider as we draw closer to the $1,000 threshold. He also devotes a chapter to his own experiences being tested by several consumer genomics companies that predict genetic risk, further illustrating both the triumphs and challenges he describes.

In *Here is a Human Being,* geneticist Misha Angrist describes his experience as the fourth person to undergo whole-genome sequencing in the Personal Genome Project. He reflects on his personal genomic journey from several perspectives—as a scientist, husband, father, son and ethicist. In this honest and deeply personal book, Angrist describes his struggles over decisions about public disclosure and privacy concerns, as well as the difficulty of making sense of the deluge of data that awaited him on the other side of the sequencer.

The Personal Genome Project offers its members whole-genome sequencing free of charge, but participants must agree to allow all of their personal information, including phenotypic characteristics, to be posted online as a demonstration of the future of medical health profiles. Angrist recalls a friend asking him, "Why in God's name would you want to do that?" In describing his participation, Angrist echoes the optimistic hopes for genomic personalized medicine expressed in the other two books under review. Yet what makes Angrist's account uniquely valuable is the degree to which it tempers the promissory language of the genomic revolution with the complexity of lived experience. "With the click of a mouse," he writes, "I had gone from an ordinary neurotic/depressive who didn't believe in genetic determinism to a quivering mess who 'knew' himself to be a ticking time bomb." Angrist's reaction provides insight into the journey facing every individual confronting such a broad spectrum of "risk information." He describes his family history of breast cancer and his worries about how his two young daughters might be affected by the public disclosure of his genome. He elucidates some of the dilemmas that will face not only patients, but also the health-care providers from whom they will seek assistance in making decisions about applying genetic information. The price tag for genome sequencing, Angrist writes, is no longer the relevant obstacle to personalized medicine; instead, interpreting what is meaningful in the vast amount of data a single human genome provides has become the ultimate challenge.

Whether recent developments in genomics fuel a revolution in personalized medicine will depend not only on scientific discoveries, but on the successful integration of genomic information into the health care system. These three books encapsulate a key moment in the development of genomic medicine, a point at which individuals, scientists, researchers and policy makers have the opportunity to shape the path forward. Understanding both the possibilities and the obstacles that these books describe will be critical in determining whether hope for a genomic future translates into more effective health care and better outcomes.

Simone Vernez is the project manager for a study of social networking and personal genomics at the Stanford Center for Biomedical Ethics. Sandra Soo-Jin Lee, who is the principal investigator for that project, is a senior research scholar at the Stanford Center for Biomedical Ethics. She is currently at work on a book titled American DNA: Race, Justice and the New Genetic Sciences.

A Bold New Bird Book

The Crossley ID Guide: Eastern Birds **(Princeton University Press, $35)** is ruffling some feathers in the birding community. It is a book that you will either love or hate. Your reaction will depend largely on whether you like to have nature presented to you in an idealized way or whether you can tolerate some of the messiness of reality. Author and photographer Richard Crossley has a high tolerance for messiness—he tries to squeeze in as much reality as he can onto every printed page.

Each of the more than 600 species in the book is shown in a lifelike scene that is typical of the species' habitat. Within that tableau, a species is represented by pictures that portray the bird at various distances, in a range of plumages and behaviors and at several ages.

Crossley crammed 10,000 photographs onto 640 plates, so, on average, each plate presents one or more species in about 15 different ways. But some birds, such as the black rail, are represented by only a single photograph, whereas others, such as the broad-winged hawk, are depicted 40 times on a single plate, as shown at right. (The 40th hawk is tricky; see if you can find it.)

It can be very difficult to identify a species from the tiny images that show the bird at a great distance, because few of the key field marks are recognizable. Why put such images in an identification guide? Crossley calls it *reality birding*. He believes that you can become a better birder by studying the distant birds and comparing them to the larger close-up images. By noticing the similarities between the different images, you will learn to focus on the features that remain constant for a particular species. The rationale is compelling, and I think Crossley's approach might actually work.

Although the plates are supposed to do most of the work in this book, Crossley does provide some text, captions and range maps to complement the images. The captions are fairly lean, and Crossley saves space by using the four-letter *alpha code* that bird banders use as shorthand for each species. Beginning birders might find the codes a little intimidating, but all of the codes are identified in an index at the back of the book.

This plate from *The Crossley ID Guide: Eastern Birds* depicts the broad-winged hawk *(Buteo platypterus*, BWHA) at various ages: 1st-s. = first summer; ad. = adult; juv. = juvenile. The plate is reproduced here at 90 percent of its actual size. The text explains that the bird, a fairly common breeder in deciduous forest, is hard to find; its high-pitched whistle is often the best clue to its presence. These hawks migrate in large flocks along the Great Lakes, down the Appalachians and through Texas. At a distance, their wings appear white with a dark border.

Much more can, and will, be said about this book. If you detest images that have been obviously manipulated using a computer, then you will almost certainly not like this book. If you prefer paintings to photographs in your guidebooks, then this book is not for you either.

For those who can tolerate Crossley's novel approach, *The Crossley ID Guide* will most likely serve as a guide that beginning and intermediate birders can study at home. The book's massive size—it is both larger and heavier than *The Sibley Guide to Birds*—and its internal organization do not lend themselves to easy use in the field. I believe that Crossley recognizes this fact, which is why the book is called an ID guide rather than a field guide.

And, in case you were wondering, I love it.—*Michael Szpir*

sigma xi today

May-June 2011 · Volume 20

Kathryn Sullivan to Receive Sigma Xi's McGovern Award

Former astronaut Kathryn D. Sullivan, the first U.S. woman to walk in space, will receive Sigma Xi's 2011 John P. McGovern Science and Society Award.

Since 1984, a highlight of Sigma Xi's annual meeting has been the McGovern Lecture, which is made by the recipient of the McGovern Medal. Recent recipients have included oceanographer Sylvia Earle and Nobel laureates Norman Borlaug, Mario Molina and Roald Hoffmann.

Sullivan was awarded a bachelor in Earth Sciences from the University of California, Santa Cruz in 1973, as well as a Ph.D. in geology from Dalhousie University in 1978.

As a member of the NASA astronaut corps, she became the first U.S. woman to walk in space in 1984. Six years later, Sullivan was part of the shuttle mission that deployed the Hubble Space Telescope.

She served as an oceanography officer in the U.S. Naval Reserve with the rank of Captain and as chief scientist for the National Oceanic and Atmospheric Administration.

After leaving NASA, Sullivan served as president and CEO of the COSI Columbus, an interactive science center in Columbus, Ohio. She currently serves as director for The Ohio State University's Battelle Center for Mathematics and Science Education Policy and as a volunteer science advisor to COSI. She was appointed to the National Science Board by President Bush in 2004.

In 2004, Sullivan was inducted into the Astronaut Hall of Fame. In 2009, Sullivan was elected to a three-year term as the Chair of the Section on General Interest in Science and Engineering for the American Association for the Advancement of Science (AAAS).

Sullivan is an adjunct professor of geology

(continued on page 272)

From the Executive Director

Annual Report

In my report last year I challenged the membership to consider the characteristics of successful associations. I suggested that we emulate what successful associations do that others do not. This year as I reflect back on the previous fiscal year, I suggest that we need to go even further.

We have intangible assets that could, if converted to tangible outcomes, add to the value of active membership in Sigma Xi. I believe that standing up for high ethical standards, encouraging the earlier career scientist and networking with colleagues of diverse disciplines is still very relevant to our professional lives. Membership in Sigma Xi still represents recognition for scientific achievements, but the value comes from sharing with companions in zealous research.

Stronger retention of members through better local programs would benefit the Society in many ways. It appears that we have continued to initiate new members in numbers similar to past years but retention has declined significantly. In addition, the source of the new members is moving more and more to the "At-large" category and less and less through the Research/Doctoral chapters.

While Sigma Xi calls itself a "chapter-based" Society, we have found that only about half of our "active" members are affiliated with chapters in "good standing." As long as chapters remain a focal point for programs, we must identify methods to revitalize and work together with the shared goals.

On the positive side, this past year has seen several new efforts and continuation of other important programs at Sigma Xi.

- Public Radio International and Sigma Xi continued our collaboration, made possible by a National Science Foundation grant. The program has been successful in launching a weekly science podcast, associated website and interactive online science discussions with expert hosts.
- The Board of Directors initiated the launch of a new networking opportunity with a Sigma Xi community on the ResearchGATE platform. ResearchGATE is a rapidly growing community of scientists and engineers in over 200 countries. The value of the network lies in the potential to encourage greater communication and research collaboration among Sigma Xi members in North America and around the world.
- The National Academy of Sciences has once again renewed its contract with Sigma Xi to distribute a special pool of NAS funds through our Grants-in-Aid of Research program. This marks the 26th year of this collaboration. Since 1922, the Society has awarded small grants to more than 30,000 undergraduate and graduate student researchers.
- The *Science in the News* e-newsletter has proven to be a popular free subscription for members and non-members. Subscribers tell us that the daily newsletter is a valuable time-saver that helps them keep up with developments in science and technology reported in the mainstream media.
- A special group of advisors met last June to discuss ethical issues related to peer review and authorship. Insights shared at the summit will help shape Sigma Xi's proposed new ethics training initiatives. We are interested in developing a companion publication to our ethical guidebooks *Honor in Science* (1984) and *The Responsible Researcher: Paths and Pitfalls* (1999). In anticipation of celebrating our 125th anniversary at the 2011 annual conference, we have selected ethics as a theme for the year.

Finally, Sigma Xi members are committed to improving the human condition through their positions as engineers and scientists. The value of membership should never be equated with the "fast-food value"— where the value of a membership lasts about as long as it takes to consume a hamburger or burrito. Let's make what we have go further by fostering integrity in research, enhancing the health of the research enterprise and promoting the public's understanding of science. Let's move forward through the remainder of 2011 with vigor and enthusiasm for our Society.

Jerome F. Baker

Sigma Xi History 1936-1961

This is the third in a series of articles about Sigma Xi's history as part of our 125th anniversary celebration.

The Society's semicentennial celebration, held at Cornell University in June 1936, attracted the attention of the *New York Times*, among other national press, through the awarding of research prizes in the physical and biological sciences. Among the young members participating in the celebration were Isidor I. Rabi and Barbara McClintock, who both later received Nobel Prizes.

With the semicentennial's success and improving economic conditions in the hope for the 1930s, Sigma Xi seemed ready for further growth. In 1937 the Executive Committee formed a Committee on Policy, whose report in 1938 suggested major changes for the *Sigma Xi Quarterly*, among other things.

The *Quarterly* had grown beyond its newsletter function and sometimes published broadly interesting articles. But it had never had an official editor or editorial board. The Society Secretary simply assembled material submitted by members and chapters.

Some members urged expansion of the *Quarterly* into a journal "within the field of science more or less equivalent to the *American Scholar* [published by the United Chapters of Phi Beta Kappa] in the field of arts and literature." But others disagreed, and the Executive Committee formed yet another committee to consider the *Quarterly's* future.

In 1939, the special committee urged a major redefinition of the *Quarterly*, recommending an increase from 200 or so pages to "some 500 to 600 pages per year." Further, while agreeing that "the *Quarterly* should continue to report" Sigma Xi news, the committee recommended an emphasis on articles presenting "recent advances in the various fields of science."

To oversee the revised journal, it recommended an editorial board and an independent (and salaried) editor in chief, authorized to commission (and pay for) "special summaries of research."

In 1941, Sigma Xi President and Yale University biologist George Baitsell, was elected Society Secretary and editor in chief of its journal. In assuming his new duties, Baitsell moved Sigma Xi's offices to Yale and focused his attention on the journal. From 1941 on, he devoted more space than ever before to articles of general scientific interest.

In 1943, G. Evelyn Hutchinson, a fellow Yale biologist, began contributing personal musings on different scientific topics to a regular feature called "Marginalia," and with his fine prose style and eye for interesting subjects he attracted many readers. "The Scientist's Book Shelf," a book review section also first appeared that year.

Emphazing science in general, the *Quarterly* in 1942 became *American Scientist*, a name first suggested in the late 1930s.

Wartime support for science brought Sigma Xi to new heights. Firms that had previously perceived research laboratories as luxuries now saw them as golden opportunities to attract federal dollars. Instrument makers, chemical suppliers and scientific publishers all advertised their products widely, and *American Scientist's* circulation rose to more than 35,000 in 1947. Other aspects of Sigma Xi also flourished under the animated leadership of President Harlow Shapley, director of the Harvard Observatory since 1921.

Sigma Xi chartered its 100th chapter in 1948, and by 1950, it boasted about 42,000 active members. In 1947, a group of Sigma Xi members formed the Scientific Research Society of America (RESA) to encourage research in government and industrial laboratories, in the same way that Sigma Xi encouraged research in the academic community.

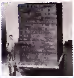
Throughout the 1950s and 1960s, Sigma Xi's growth continued until it had more than doubled in size.

American Scientist,
September-October 1986

1936

Sigma Xi celebrates 50 years

50



1942

Enrico Fermi (SX 1939) assembled the first artificial nuclear reactor

Sigma Xi Quarterly* becomes *American Scientist

J. Robert Oppenheimer (SX 1928) scientific director of the "Manhattan Project"

1945

Vannevar Bush's (SX 1934) "Science—the Endless Frontier" letter appears

1947

Scientific Research Society of America (RESA) founded
(chapters later became Sigma Xi chapters)

John Bardeen (SX 1929), Walter Brattain (SX 1925) and William Shockley (SX 1932) invent the first transistor

New Sigma Xi Chapters • 1936-1962

Carnegie-Mellon Univ.
George Washington Univ.
Oregon State Univ.
Univ. of Utah
Rice Univ.-TX Med. Ctr.
Univ. of Florida
Univ. of MA
Wellesley College
Univ. of Alabama
West Virginia Univ.
Univ. of Southern CA
Virginia Tech Univ.
Bryn Mawr College
Oberlin College
Illinois Inst. of Tech.
IL State Univ./IL Wesleyan Univ.
Louisiana State Univ.
Utah State Univ.
Polytechnic/SUNY Old Westbury
Tufts Univ.
Depauw-Wabash
Emory Univ.
NC State Univ.
Saint Louis Univ.
Vanderbilt Univ.
Wayne State Univ.
Abbott Laboratories
Catholic Univ. of America
Univ. of TX SW Med. Ctr.
Univ. of Connecticut
Corning Inc.
Univ. of Georgia
Vermont
Lynchburg
Socony Mobil
Rockford College
Univ. of CA-Davis
Univ. of Hawaii
San Diego
TN A & I Fisk-Meharry
Univ. of Maine
Frederick
Oklahoma State Univ.
Roche Research ■
Temple Univ.

Amherst College
Auburn Univ.
Brigham Young Univ.
Butler-Indianapolis ■
CA State Univ.-Fresno
CDC ■
China Lake ■
Montana State Univ.
Naval Research Lab.-Edison ■
Univ. of Denver
Univ. of Tennessee
Albany ■
G. E. Electronics Lab. ■◆
NY Univ. College of Dentistry ■
Texas A & M Univ.
Alamo
Alfred Univ.
Argonne National Lab. ■
Boston Univ.
Columbia-Willamette
DuPont ■
Franklin Research Ctr. ■◆
Hanscom
Kansas City
Texaco Research Ctr. ■◆
Triple Cities New York
Univ. of Notre Dame
Alcoa ■
Colorado State Univ.
Georgia Inst. of Tech.
GTE Labs. ■
M.W. Kellogg ■
Beloit College ●
Rollins College
San Jose State Univ.
Tuskegee Univ.
Univ. of AK-Fayetteville
Univ. of RI
Brooklyn College
Franklin & Marshall College
Lamar Univ.
Univ. of Redlands ◆
Lever Research Ctr. ◆
MO Univ. of S&T
Oak Ridge ■
Stevens Inst. of Tech.

Univ. of NH
Univ. of NM
Univ. of TX Med. Branch
Atlanta Univ. Ctr.
CA Polytechnic State Univ.
Florida State Univ.
Hughes Labs. ■
Louisville
Univ. of LA at Lafayette
Univ. of North Texas
Boston College ◆
Humboldt State Univ.
Loyola Univ.
Texas Womans Univ. ◆
Univ. of MN at Duluth
Univ. of TN Memphis
Alaska
American Univ. of Beirut
Army Research Lab. (ARL)
Chesapeake ■
Denison Univ.
Howard Univ.
Natick ■
New Mexico State Univ.
New Orleans ■
Ohio Wesleyan Univ.
Southern Univ.
Univ. of AL-Birmingham
Univ. of South Dakota
CA State Univ.-Chico
Fordham Univ.
Drake Univ. ▲
Gulf Coast
Midland ■◆
Monsanto Science ■◆
Naval Postgraduate School
Rockefeller Univ.
Saint John's Univ. ◆
Tahoma
Univ. of SC
Adelphi Univ. ◆
Villanova Univ.
Andrews-Whirlpool ■
Knox College
Nalco Chemical ■
Northern Westchester ■

Southern Indiana
Univ. of Delaware
Vassar College
Ventura County ■
Wooster Ohio
Albion College
Claremont Colleges
Hollins Univ.
Jacksonville
Marquette Univ.
NM Highlands Univ.
Olin ▲
Portland State Univ. ◆
San Fran. State Univ.

Texas Tech Univ.
Univ. of MA-Lowell
Boeing NA
Eckerd College
Ford Motor Company ■
Lafayette College
Ohio Univ.
South Dakota State Univ.
Univ. of New Orleans ■
Univ. of PR at Mayaguez
Ball State Univ.
CA State Univ.-Hayward
Central Michigan
Hartford ■

LCA Research ■◆
Rosalind Franklin Univ.
SUNY at Stony Brook
Univ. of LA at Monroe
Tidewater Virginia
Univ. of The South
Whitman College-Walla Walla Univ.

■ *RESA Chapter*
◆ *Disbanded*
▲ *Now part of Greater New Orleans*
● *Now part of Univ. of WI*

Sigma Xi Nobel Laureates

Chemistry
1946 John H. Northrop
1946 James B. Sumner
1948 Arne Tiselius
1949 William F. Giauque
1951 Edwin M. McMillan
1951 Glenn T. Seaborg
1954 Linus Pauling
1955 Vincent du Vigneaud
1959 Jaroslav Heyrovsky
1960 Willard F. Libby
1961 Melvin Calvin

Peace
1962 Linus Pauling

Physics
1937 Clinton Davisson
1938 Enrico Fermi
1939 Ernest Lawrence
1943 Otto Stern
1944 Isidor Isaac Rabi
1945 Wolfgang Pauli
1946 Percy W. Bridgman
1952 Felix Bloch
1952 E. M. Purcell
1955 Polykarp Kusch
1955 Willis E. Lamb
1956 John Bardeen
1956 Walter H. Brattain
1956 William B. Shockley
1957 Chen Ning Yang
1959 Owen Chamberlain
1959 Emilio G. Segrè
1960 Donald A. Glaser
1961 Robert Hofstadter

Physiology or Medicine
1943 Henrik Dam
1943 Edward A. Doisy
1944 Joseph Erlanger
1944 Herbert S. Gasser
1946 Hermann J. Muller
1947 Carl Cori
1947 Gerty Cori
1950 Philip S. Hench
1950 Edward C. Kendall
1951 Max Theiler
1952 Selman A. Waksman
1953 Fritz Lipmann
1954 John F. Enders
1954 Frederick C. Robbins
1954 Thomas H. Weller
1956 Dickinson W. Richards
1958 George Beadle
1958 Joshua Lederberg
1958 Edward Tatum
1959 Arthur Kornberg
1961 Georg von Békésy
1962 Francis Crick
1962 James Watson

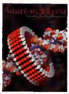
1948
100th Sigma Xi chapter was chartered

1950
The National Science Foundation is created

1952
Grace Hopper (SX1934) developed the first compiler for a computer language

Jonas Salk (SX 1945) developed polio vaccine

1953
Francis Crick (SX 1954) and James Watson (SX 1959) discovered the structure of DNA



1957
Amundsen-Scott South Pole Station established

1960
The first communications satellite, Echo 1, is launched

McGovern Award

(continued on page 296)

at The Ohio State University and a Fellow of the AAAS and the American Institute of Astronautics and Aeronautics. A Sigma Xi member since 1989, she is also a member of the Woods Hole Oceanographic Institution, the Explorers Club, the Society of Woman Geographers and Association of Space Explorers.

She has been awarded honorary degrees by five universities. Her honors include the *Aviation Week & Space Technology* Aerospace Legend Award; Astronaut Hall of Fame; and the Public Service Award from the National Science Board, in recognition of lifelong commitment to science education.

She has also received the Juliette Award for National Women of Distinction, Girl Scouts USA; the Lone Sailor Award, U.S. Navy Memorial Foundation; NASA Medal for Outstanding Leadership; AIAA Haley Space Flight Award; AAS Space Flight Achievement Award; and the National Air and Space Museum Trophy, Smithsonian Institution.

In January 2011, Sullivan was nominated by President Barack Obama to be an assistant secretary of commerce.

U.S. Rep. David Price Named Honorary Sigma Xi Member

U.S. Congressman David Price (D-North Carolina) will be inducted as an honorary member of Sigma Xi at the Society's annual meeting next November in Raleigh, North Carolina.

Beginning in 1983, distinguished individuals not otherwise eligible for membership in Sigma Xi, who have served science, or the Society, in a manner or to a degree that merits such recognition, have been elected honorary life members by the Board of Directors.

Price was named a "Champion of Science" by the Science Coalition, a non-profit, nonpartisan organization of 50 of the leading public and private research universities in the U.S. The North Carolina chapter of the Sierra Club has also recognized him as their "Legislator of the Year."

Among other accomplishments, Price authored a bill to establish the National Science Foundation's Advanced Technological Education program, which helps community colleges upgrade their training programs for jobs in high-tech fields. From his position on the Appropriations Committee, he supported the effort to double funding for the National Institutes of Health over five years and has pressed for a similar boost to the National Science Foundation.

He is a recipient of the American Political Science Association's Hubert H. Humphrey Public Service Award and also has been recognized by Voices for AmeriCorps, the NC Low-Income Housing Coalition, the NC Public Transportation Association, the Association of Community College Trustees, the National Association of Land-Grant Colleges and State Universities, and other organizations for his work in housing, education, and transportation policy.

Before going to Congress in 1987, Price was a professor of political science and public policy at Duke University. He is the author of four books on Congress and the American political system.

Born in 1940, Price grew up in the small town of Erwin in eastern Tennessee. His father was a high school principal, and his mother was an English teacher.

Price was a Morehead Scholar at the University of North Carolina at Chapel Hill. He earned his B.A. in 1961 and continued his education at Yale University, where he received a Bachelor of Divinity degree and a Ph.D. in political science.

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL
FIRST-CLASS MAIL PERMIT NO. 1348 DURHAM NC

POSTAGE WILL BE PAID BY ADDRESSEE

AMERICAN
Scientist
P.O. Box 13975
Research Triangle Park, NC 27709-9890

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL
FIRST-CLASS MAIL PERMIT NO. 1348 DURHAM NC

POSTAGE WILL BE PAID BY ADDRESSEE

AMERICAN
Scientist
P.O. Box 13975
Research Triangle Park, NC 27709-9890

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL
FIRST-CLASS MAIL PERMIT NO. 1348 DURHAM NC

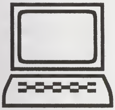
POSTAGE WILL BE PAID BY ADDRESSEE

SIGMA XI
THE SCIENTIFIC RESEARCH SOCIETY
P.O. Box 13975
Research Triangle Park, NC 27709-9890

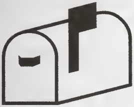
AMERICAN
Scientist

4 Ways to Subscribe:

Online
www.americanscientist.org



Mail
Attached card



Call
800-282-0444 or
919-549-0097

Fax
919-595-8114